

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 22
2016
№ 8

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

УЧРЕДИТЕЛЬ
Издательство "Новые технологии"

СОДЕРЖАНИЕ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

- Черняховская Л. Р., Галиуллина А. Ф., Сабитов И. И.** Оценка эффективности предоставления государственных услуг с использованием системы нечеткого логического вывода 563
- Зак Ю. А.** Алгоритмы решения транспортной задачи в условиях нечетких данных о стоимости поставки грузов 568
- Бутаков С. В., Мурзинцев С. В., Цхай А. А.** Использование горизонтально масштабируемой инфраструктуры при поиске заимствований в тексте. 576

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ

- Левин В. И.** Модели производной и их применение для моделирования динамических неполностью определенных систем 582
- Исаев Д. В.** Имитационное моделирование сложных проектов. 589
- Моисеев А. А.** Критериальное моделирование в виртуализации квалификационных испытаний 596
- Доморацкий Е. П.** Алгоритм синтеза геометрических характеристик проекционных изображений трехмерных объектов 605

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

- Саак А. Э.** О качестве диспетчеризации массивов точных форм с заявками параболического типа в Grid-системах 610

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ

- Димов Э. М., Маслов О. Н., Раков А. С.** Управление информационной безопасностью корпорации с применением критериев риска и ожидаемой полезности 620
- Еременко А. В., Сулавко А. Е., Толкачева Е. В., Левитская Е. А.** Метод защиты текстовых документов на электронных и бумажных носителях на основе скрытого биометрического идентификатора субъекта, получаемого из подписи 628

БАЗЫ ДАННЫХ

- Чернышов А. В.** К вопросу о применении оптических дисков для создания долговременных электронных архивных хранилищ информации небольших организаций 635

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.
МИХАЙЛОВ Б. М., д. т. н., проф.
НЕЧАЕВ В. В., к. т. н., проф.
ПОЛЕЩУК О. М., д. т. н., проф.
САКСОНОВ Е. А., д. т. н., проф.
СОКОЛОВ В. 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

INTELLIGENT SYSTEMS AND TECHNOLOGIES

- Chernyakhovskaya L. R., Galiullina A. F., Sabitov I. I.** Efficiency Evaluation of Public Services Provision Using Fuzzy Inference System 563
- Zack Yu. A.** Algorithms for Solving the Transport Problem in Fuzzy Data on the Cost of Delivery of Goods 568
- Butakov S. V., Murzintsev S. V., Tskhai A. A.** Detecting Text Similarity on a Scalable Cluster 576

MODELING AND OPTIMIZATION

- Levin V. I.** Models of Derivative and their Application to Modeling of Dynamical Incompletely Defined Systems 582
- Isaev D. V.** Simulation of Complicated Projects 589
- Moiseev A. A.** Criteria Modeling in Virtualization of Qualification Tests 596
- Domorackiy E. P.** Synthesis Algorithm of the Geometrical Characteristics of the Projection Images of Three-Dimensional Objects 605

COMPUTING SYSTEMS AND NETWORKS

- Saak A. E.** To Quality of Scheduling of Precise Form Sets of Parabolic Type Tasks in Grid Systems 610

CRYPTOSAFETY INFORMATION

- Dimov Je. M., Maslov O. N., Rakov A. S.** Information Security Management Corporation of Corporation Using the Risk Criterion and Expected Utility 620
- Eremenko A. V., Sulavko A. E., Tolkacheva E. V., Levitskaya E. A.** The Method of Protection of Electronic and Paper-Based Text Documents with Biometric Identifier of the Subject Obtained from his Signature 628

DATABASE

- Chernyshov A. V.** To the Question of the Optical Discs Application for Long Term Digital Archive Storage of Small Organizations. 635

Editor-in-Chief:

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

Deputy Editor-in-Chief:

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

Chairman:

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

Editorial Board Members:

Avdoshin S. M., Cand. Sci. (Tech.), Ass. Prof.
Antonov B. I.
Barsky A. B., Dr. Sci. (Tech.), Prof.
Vasenin V. A., Dr. Sci. (Phys.-Math.), Prof.
Vishnekov A. V., Dr. Sci. (Tech.), Prof.
Galushkin A. I., Dr. Sci. (Tech.), Prof.
Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.
Domrachev V. G., Dr. Sci. (Tech.), Prof.
Zaborovsky V. S., Dr. Sci. (Tech.), Prof.
Zagidullin R. Sh., Cand. Sci. (Tech.), Ass. Prof.
Zarubin V. S., Dr. Sci. (Tech.), Prof.
Karpenko A. P., Dr. Sci. (Phys.-Math.), Prof.
Kolin K. K., Dr. Sci. (Tech.)
Kulagin V. P., Dr. Sci. (Tech.), Prof.
Kureichik V. M., Dr. Sci. (Tech.), Prof.
Ljovitch Ya. E., Dr. Sci. (Tech.), Prof.
Mikhailov B. M., Dr. Sci. (Tech.), Prof.
Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.
Poleschuk O. M., Dr. Sci. (Tech.), Prof.
Saksonov E. A., 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"

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ INTELLIGENT SYSTEMS AND TECHNOLOGIES

УДК 004.8:(351/354)

Л. Р. Черняховская, д-р техн. наук, проф., e-mail: lrchern@yandex.ru,

А. Ф. Галиуллина, аспирант, e-mail: GAF1205@yandex.ru,

И. И. Сабитов, аспирант, e-mail: iskra1990@gmail.com,

ФГБОУ ВПО "Уфимский государственный авиационный технический университет" (УГАТУ)

Оценка эффективности предоставления государственных услуг с использованием системы нечеткого логического вывода

Представлена модель оценки эффективности системы предоставления государственных услуг, рассматриваемой как система массового обслуживания, разработанная в программном комплексе MATLAB/Simulink. Данная модель использует систему нечеткого логического вывода и при низкой эффективности системы позволяет оценить влияние управляющего воздействия в обратной связи на эффективность системы массового обслуживания посредством изменения времени обслуживания и затрат ресурсов на обслуживание.

Ключевые слова: система предоставления государственных услуг, система массового обслуживания, оценка эффективности, система нечеткого логического вывода

Введение

Ряд задач повышения эффективности предоставления государственных услуг связан с совершенствованием систем массового обслуживания (СМО), т. е. таких систем, в которых, с одной стороны, возникают массовые запросы на выполнение каких-либо услуг, а с другой — происходит удовлетворение этих запросов [1–3].

Модель СМО включает в себя такие элементы, как источник требований, генерирующий входящий поток требований, очередь, обслуживающие устройства (каналы обслуживания), выходящий поток требований. Исследованием таких систем занимается теория массового обслуживания.

Методами теории массового обслуживания могут быть решены многие задачи исследования процессов, происходящих в системе государственного управления. Так, в организации системы предоставления государственных услуг эти методы позволяют определить оптимальное число многофункциональных центров, численность специалистов, непосредственно предоставляющих гражданам государственные услуги, и другие параметры. При этом необходимо добиваться того, чтобы суммарные расходы на обслуживание и убытки от простоя были минимальными, а эффективность массового обслуживания — максимальной.

Обеспечение качества государственных услуг требует от системы предоставления государственных услуг повышенной надежности, эффективности функционирования, а также гибкости в ответ на динамические изменения внешней среды, снижения

рисков и потерь от запоздалых и некомпетентных управленческих решений.

Необходимо оперативно отслеживать изменения в показателях эффективности СМО и принимать компетентные управленческие решения в зависимости от складывающихся проблемных ситуаций, связанных главным образом со снижением эффективности массового обслуживания.

Согласно ГОСТ Р ИСО 9000—2008 "Системы менеджмента качества. Основные положения и словарь", **эффективность** — это связь между достигнутым результатом и использованными ресурсами [4].

Эффективность функционирования СМО определяется различными показателями: абсолютной пропускной способностью, относительной пропускной способностью, временем обслуживания заявок, временем ожидания, средним числом занятых каналов и др. Некоторые из этих показателей могут использоваться в качестве критериев оценки эффективности СМО. Более подробно показатели эффективности одноканальной СМО с отказами, многофазной многоканальной СМО с ожиданием исследованы в работах [5–8], где также проведена их оценка.

При решении задач, связанных с массовым обслуживанием, большое значение имеет правильный выбор критериев, определяющих изучаемый процесс. Одна и та же система обслуживания может характеризоваться с различных точек зрения различными критериями эффективности. Выбор того или иного критерия должен проводиться в каждом

конкретном случае исходя из тех задач, которые ставятся перед системой.

На практике в системе предоставления государственных услуг используются многоканальные СМО с ожиданием, для оценки эффективности которых будем рассматривать три наиболее важных, с нашей точки зрения, критерия: время обслуживания, качество обслуживания и затраты ресурсов на обслуживание.

Для обеспечения требуемой эффективности системы предоставления государственных услуг необходимо установление зависимостей эффективности СМО от ее организации, которая обусловлена установкой упомянутых выше параметров СМО.

Разработка системы нечеткого логического вывода для определения эффективности системы массового обслуживания

На деятельность системы предоставления государственных услуг влияют неопределенные факторы. К ним можно отнести факторы внешней и внутренней среды, значения которых неизвестны или известны не полностью. Несмотря на дефицит информации необходимо учитывать возможное влияние неопределенных факторов на систему предоставления государственных услуг.

Учет фактора неопределенности при решении задач, связанных с массовым обслуживанием, во многом изменяет методы принятия решения: меняется принцип представления исходных данных и параметров модели, становятся неоднозначными понятия решения задачи и оптимальности решения. Поэтому предлагается разработать систему нечеткого логического вывода с применением теории нечетких множеств.

В практике управления сложными динамическими системами, к которым также можно отнести и систему предоставления государственных услуг, обычно не придерживаются границ области допустимых значений контролируемых параметров, и многие ограничения чаще всего являются "мягкими", допускающими их незначительное нарушение. Границы области критических режимов также могут быть нечеткими. Применение нечетких ("мягких") ограничений значительно расширяет возможность контроля и управления и делает их адекватными реальной обстановке в системе [9], а также предоставляет возможность применения экспертных знаний, выраженных с использованием лингвистических переменных.

Система нечеткого логического вывода позволяет получать заключения о требованиях, предъявляемых к управлению какой-либо системой на основе нечетких условий или предпосылок, представляющих собой информацию о допустимых характеристиках состояний этой системы.

Для проведения нечеткого моделирования существуют различные инструментальные средства, например, *FuzzyTECH*, *CubiCalc*, *MATLAB*. В рассматриваемых исследованиях нечеткое моделиро-

вание проведено в среде *MATLAB* с помощью пакета расширения *Fuzzy Logic Toolbox*, в котором реализовано множество функций нечеткой логики и нечеткого вывода. *Fuzzy Logic Toolbox* обладает простым и хорошо продуманным интерфейсом, позволяющим легко проектировать и диагностировать нечеткие модели. Графические средства *Fuzzy Logic Toolbox* позволяют интерактивно отслеживать особенности поведения системы.

Для определения эффективности СМО нами в редакторе *FIS* определены три входных лингвистических переменных: *ServiceTime* (время обслуживания), *ServiseQuality* (качество обслуживания), *ResourceCost* (затраты ресурсов на обслуживание) и одна выходная переменная — *Efficiency* (эффективность СМО).

Описание входных и выходной переменных представлено в табл. 1 и 2 соответственно.

Таблица 1

Описание входных переменных

Наименование	Диапазон изменения	Термы	Тип функций принадлежности
<i>ServiceTime</i>	0...1	<i>VS</i> — <i>very short</i> (очень короткое) <i>S</i> — <i>short</i> (короткое) <i>M</i> — <i>middle</i> (среднее) <i>L</i> — <i>long</i> (длительное) <i>VL</i> — <i>very long</i> (очень длительное)	Гауссова
<i>ServiseQuality</i>	0...1	<i>VL</i> — <i>very low</i> (очень низкое) <i>L</i> — <i>low</i> (низкое) <i>M</i> — <i>middle</i> (среднее) <i>H</i> — <i>high</i> (высокое) <i>VH</i> — <i>very high</i> (очень высокое)	Гауссова
<i>ResourceCost</i>	0...1	<i>VS</i> — <i>very small</i> (очень маленькие) <i>S</i> — <i>small</i> (маленькие) <i>M</i> — <i>middle</i> (средние) <i>B</i> — <i>big</i> (большие) <i>VB</i> — <i>very big</i> (очень большие)	Гауссова

Таблица 2

Описание выходной переменной

Наименование	Диапазон изменения	Термы	Значение
<i>Efficiency</i>	0...1	<i>LowEfficiency</i> (низкая эффективность)	0...0,3
		<i>AverageEfficiency</i> (средняя эффективность)	0,3...0,6
		<i>HighEfficiency</i> (высокая эффективность)	0,6...1

Таблица 3

Правила построения системы нечеткого логического вывода

№	ЕСЛИ			ТО
	<i>ServiceTime</i>	<i>ServiceQuality</i>	<i>ResourceCost</i>	<i>Effeciency</i>
1	<i>VS</i>	<i>VL</i>	<i>VS</i>	<i>LowEffeciency</i>
2	<i>VS</i>	<i>VL</i>	<i>S</i>	<i>LowEffeciency</i>
3	<i>VS</i>	<i>VL</i>	<i>M</i>	<i>LowEffeciency</i>
4	<i>VS</i>	<i>VL</i>	<i>B</i>	<i>LowEffeciency</i>
5	<i>VS</i>	<i>VL</i>	<i>VB</i>	<i>LowEffeciency</i>
6	<i>VS</i>	<i>L</i>	<i>VS</i>	<i>LowEffeciency</i>
7	<i>VS</i>	<i>L</i>	<i>S</i>	<i>LowEffeciency</i>
8	<i>VS</i>	<i>L</i>	<i>M</i>	<i>LowEffeciency</i>
9	<i>VS</i>	<i>L</i>	<i>B</i>	<i>LowEffeciency</i>
10	<i>VS</i>	<i>L</i>	<i>VB</i>	<i>LowEffeciency</i>
11	<i>VS</i>	<i>M</i>	<i>VS</i>	<i>HighEffeciency</i>
12	<i>VS</i>	<i>M</i>	<i>S</i>	<i>HighEffeciency</i>
13	<i>VS</i>	<i>M</i>	<i>M</i>	<i>HighEffeciency</i>
14	<i>VS</i>	<i>M</i>	<i>B</i>	<i>AverageEffeciency</i>
15	<i>VS</i>	<i>M</i>	<i>VB</i>	<i>LowEffeciency</i>
...
121	<i>VL</i>	<i>VH</i>	<i>VS</i>	<i>AverageEffeciency</i>
122	<i>VL</i>	<i>VH</i>	<i>S</i>	<i>AverageEffeciency</i>
123	<i>VL</i>	<i>VH</i>	<i>M</i>	<i>LowEffeciency</i>
124	<i>VL</i>	<i>VH</i>	<i>B</i>	<i>LowEffeciency</i>
125	<i>VL</i>	<i>VH</i>	<i>VB</i>	<i>LowEffeciency</i>

С помощью модуля *fuzzy* в *MATLAB* задаются продукционные правила для разрабатываемой системы нечеткого логического вывода. Продукционные правила — это условные высказывания типа "ЕСЛИ — ТО", которые представляют собой нечеткую импликацию:

ЕСЛИ x_1 есть A_1 И x_2 есть A_2 И ... x_n есть A_n
ТО y есть B .

Здесь A_i , B — это лингвистические значения, идентифицированные нечетким способом через соответствующие функции принадлежности $\mu A(x_i)$ и $\mu B(y)$ для переменных x и y .

При формировании базы правил необходимо учитывать весовой коэффициент каждого правила,

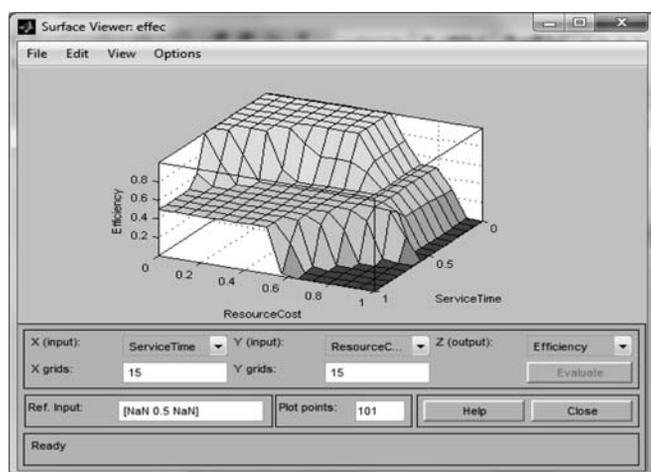


Рис. 1. Визуализация поверхности нечеткого вывода рассматриваемой модели для входных переменных "Время обслуживания" и "Затраты ресурсов на обслуживание"

который выражает количественную оценку степени истинности или относительный вес нечеткого правила. Весовой коэффициент принимает свое значение из интервала от 0 до 1 и устанавливается для каждого правила. В табл. 3 представлены правила, сформулированные на основе экспертных знаний.

Выполним анализ построенной системы нечеткого логического вывода для рассматриваемой задачи определения эффективности системы предоставления государственных услуг. Анализ системы выполним с помощью программы просмотра правил, которая предоставляет возможность получить значения выходных переменных в зависимости от значений входных.

Так, для частного случая, когда значения входной переменной "время обслуживания" равно 0,236; "качество обслуживания" — 0,801; "затраты ресурсов на обслуживание" — 0,523, процедура нечеткого вывода, выполненная системой *MATLAB* для разработанной нечеткой модели, выдает в результате значение выходной переменной "Эффективность", равное единице, что соответствует высокой эффективности при коротком времени обслуживания, высоком уровне качества обслуживания и средних затратах ресурсов на обслуживание.

Для общего анализа разработанной нечеткой модели рассмотрим визуализацию соответствующей поверхности нечеткого вывода, которая позволяет проанализировать зависимость выходной переменной от отдельных входных переменных.

По трехмерной поверхности нечеткого вывода для входных переменных "Время обслуживания" и "Затраты ресурсов на обслуживание" (рис. 1) можно сказать, что при минимальных значениях данных переменных значение выходной переменной равно 1, что соответствует высокой эффективности.

По следующей поверхности (рис. 2), отражающей зависимость выходной переменной от качества обслуживания и времени обслуживания, видно,

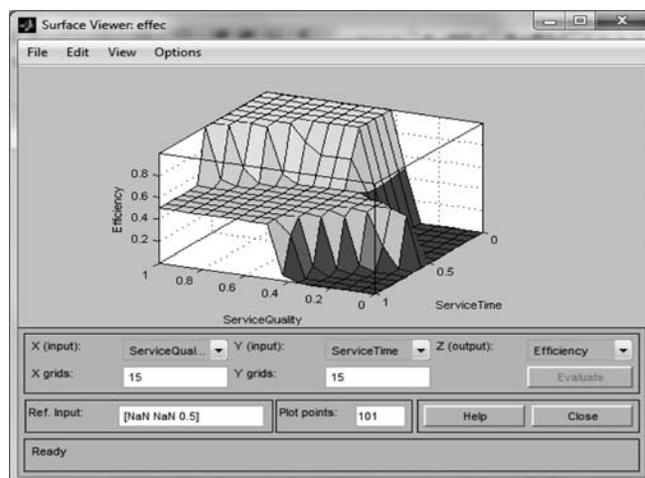


Рис. 2. Визуализация поверхности нечеткого вывода рассматриваемой модели для входных переменных "Качество обслуживания" и "Время обслуживания"

что с увеличением качества обслуживания и уменьшением времени обслуживания эффективность можно отнести к высокой.

Полученные результаты свидетельствуют о том, что разработанная база знаний адекватна и отвечает представлениям экспертов об эффективности системы предоставления государственных услуг.

Оценка эффективности системы массового обслуживания в среде *Simulink/MATLAB*

Нами была разработана модель, позволяющая провести оценку эффективности СМО на основе разработанной выше системы нечеткого логического вывода (рис. 3).

Модель включает в себя следующие компоненты и моделирующие их блоки:

- три источника случайного сигнала (*Random Source*), генерирующие случайным образом значения входных переменных (показателей эффективности) — времени обслуживания (*ServiceTime*), качества обслуживания (*ServiceQuality*) и затрат ресурсов на обслуживание (*ResourceCost*), — в интервале от 0,1 до 1;
- три нечетких контроллера для включения в модель разработанной системы нечеткого логического вывода (*Fuzzy Logic Controller*);
- два блока визуализации процесса моделирования (*Signal Scope*);
- семь мультиплексоров, которые объединяют входные сигналы в вектор (*Mux*);
- два компаратора, необходимых для сравнения эффективности СМО с заданным значением 0,3 (*Compare To Constant*);

- два усилителя, выполняющих умножение входного сигнала на постоянный коэффициент 0,1 (*Gain*);
- два блока умножения, выполняющих вычисление произведения текущих значений сигналов (*Product*);
- два блока передаточной функции (*Transfer Fcn*);
- два блока вычисления суммы (*Sum*);
- два источника постоянного сигнала, задающих постоянный по уровню сигнал (*Constant*).

Разработанная модель содержит две обратные связи, которые предназначены для осуществления управляющего воздействия на время обслуживания и затраты ресурсов на обслуживание. Модель при эффективности ниже 0,3 (ниже 30 %), т.е. при низкой эффективности СМО, поэтапно уменьшает значения двух входных переменных на 10 %: времени обслуживания и затрат ресурсов на обслуживание, — в целях повышения эффективности СМО.

Задав время моделирования — 40 и запустив моделирование, по окончании получаем результаты, представленные на рис. 4 (эксперимент 1) и рис. 5 (эксперимент 2). На данных рисунках слева представлены входные переменные (время обслуживания, качество обслуживания и затраты ресурсов на обслуживание), справа — выходная переменная (эффективность).

Как видно из рис. 4, в интервале времени [0; 5] эффективность СМО была очень низкой, менее 0,3, поэтому время обслуживания уменьшили на 10 % и довели от 0,6 до 0,54. Это помогло повысить эффективность до почти 0,2, однако желаемого результата достичь не удалось (эффективность оставалась меньше 0,3). Затем были уменьшены затраты ресурсов на обслуживание на 10 % (от 0,75 до 0,68),

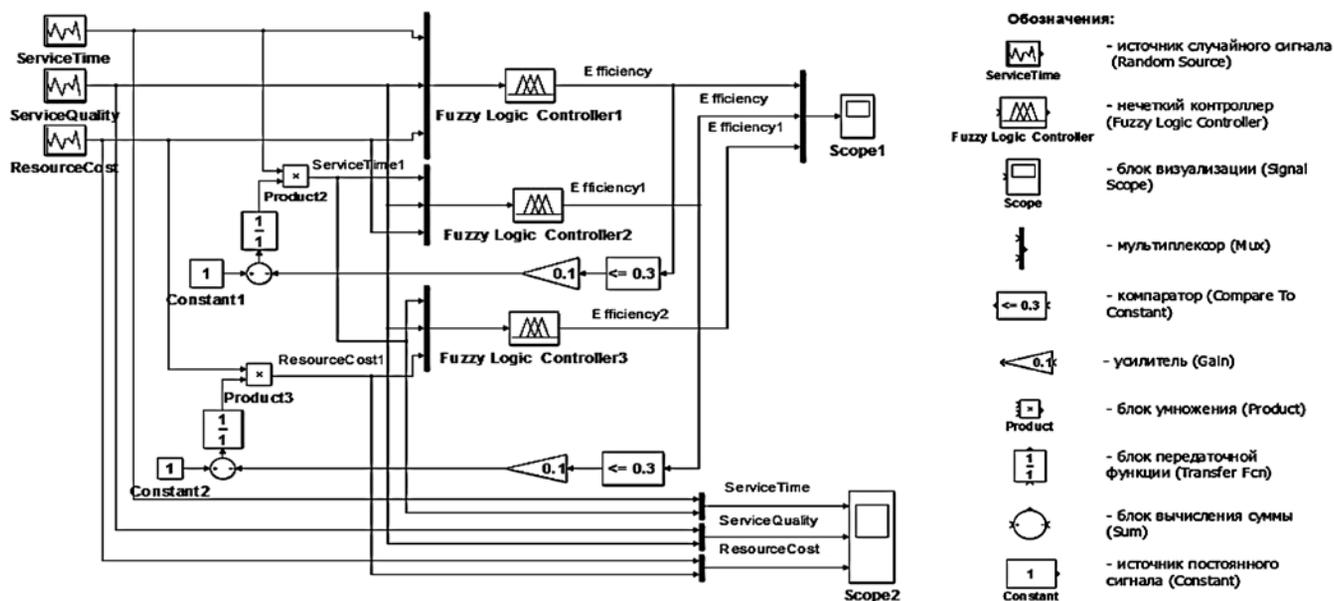


Рис. 3. Схема модели оценки эффективности СМО

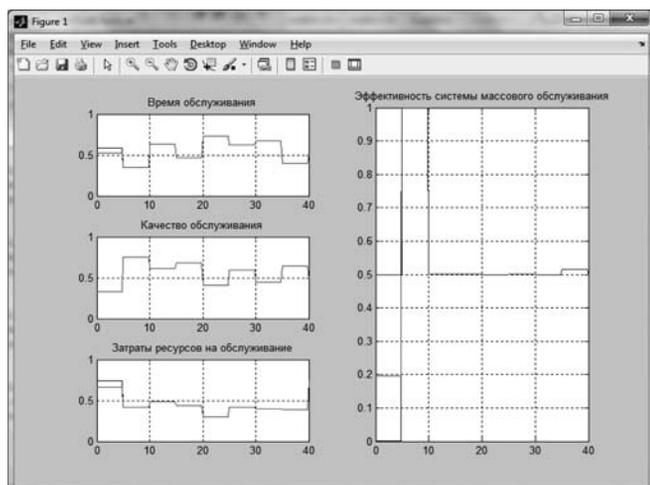


Рис. 4. Результаты моделирования (эксперимент 1)

что позволило повысить эффективность СМО до 0,5, что выше 0,3, т. е. эффективность средняя.

Как видно из рис. 5, в двух случаях, в интервале времени [15; 20] и [20; 25], эффективность СМО была низкой (менее 0,3), поэтому время обслуживания уменьшили на 10 %. В первом случае в интервале времени [15; 20] время обслуживания снизили с 0,62 до 0,56. Это помогло повысить эффективность с 0,25 до 0,5. Во втором случае в интервале времени [20; 25] время обслуживания снизили с 0,74 до 0,67. Это помогло повысить эффективность с 0,3 до 0,5. Так как в обоих случаях после уменьшения времени обслуживания был достигнут желаемый результат, эффективность стала средней (более 0,3), то уменьшение затрат ресурсов на обслуживание не понадобилось.

Таким образом, разработанная модель позволяет оценить влияние управляющего воздействия в обратной связи на эффективность системы массового обслуживания посредством изменения значений таких параметров системы, как время обслуживания и затраты ресурсов на обслуживание. Результаты моделирования показали, что управляющее воздействие оказывает положительное влияние на эффективность системы и позволяет повысить ее до требуемого уровня.

Заключение

В данной статье представлена разработанная нами система нечеткого логического вывода для определения эффективности системы массового обслуживания, которая может быть использована в системе предоставления государственных услуг. На основе продукционных правил данная система позволяет сделать вывод об эффективности СМО. Кроме того, разработана модель оценки эффективности СМО, основанная на системе нечеткого логического вывода и позволяющая при низкой эффективности

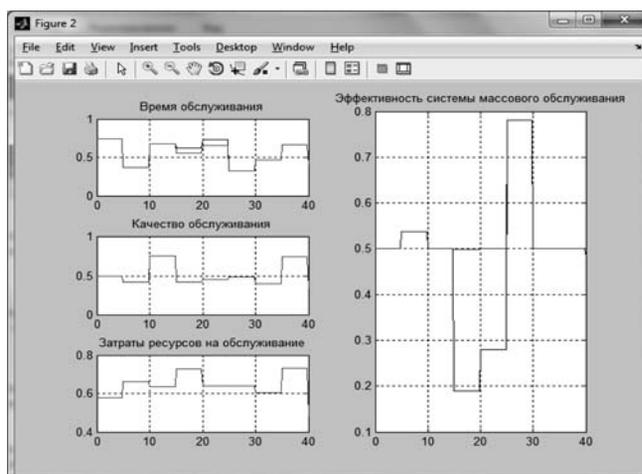


Рис. 5. Результаты моделирования (эксперимент 2)

системы провести оценку влияния управляющего воздействия в обратной связи на эффективность СМО посредством уменьшения времени обслуживания и затрат ресурсов на обслуживание. Результаты моделирования показали положительное влияние управляющего воздействия на эффективность СМО.

Работа выполнена в соответствии с грантом РФФИ 14-08-97023 "Интеллектуальная поддержка принятия решений при управлении инновационными проектами на основе обработки знаний и математического моделирования".

Список литературы

- Советов Б. Я., Яковлев С. А. Моделирование систем: учеб. для вузов. 6-е изд. стер. М.: Высш. школа, 2009. 344 с.
- Хинчин А. Я. Работы по математической теории массового обслуживания / Под ред. Б. В. Гнеденко. 3-е изд. М.: Эдиториал УРСС, 2009. 240 с.
- Черняховская Л. Р. Моделирование и анализ характеристик динамических процессов в условиях неопределенности с применением теории массового обслуживания: учеб. пособие. Уфа: УГАТУ, 2014. 159 с.
- ГОСТ Р ИСО 9000—2008. Системы менеджмента качества. Основные положения и словарь. Национальный стандарт Российской Федерации. М.: Стандартинформ, 2009. 35 с.
- Галиуллина А. Ф. Имитационное моделирование деловых процессов как систем массового обслуживания // Актуальные проблемы науки и техники. Девятая Всероссийская зимняя школа-семинар аспирантов и молодых ученых. Том 1. Информационные и инфокоммуникационные технологии. Сб. трудов. Уфа: УГАТУ, 2014. С. 226—230.
- Галиуллина А. Ф., Сильнова С. В., Черняховская Л. Р. Оценка эффективности управления производственным процессом с применением имитационного моделирования на основе систем массового обслуживания // Вестник УГАТУ. Уфа: УГАТУ, 2015. Т. 19, № 1 (67). С. 115—122.
- Галиуллина А. Ф. Аналитическое моделирование деловых процессов как систем массового обслуживания для оценки их показателей эффективности // Наука вчера, сегодня, завтра. Сб. статей по материалам XX междунар. науч.-практ. конф. № 1 (18). Новосибирск: Изд. "СибАК", 2015. С. 11—15.
- Галиуллина А. Ф. Моделирование процесса предоставления государственных услуг // Научное обозрение. 2015. № 22. С. 253—263.
- Леоненков А. В. Нечеткое моделирование в среде Matlab и fuzzyTECH. СПб.: БХВ-Петербург, 2003. 736 с.

L. R. Chernyakhovskaya, Dr. of Tech. Sci., Prof., lrchern@yandex.ru,
A. F. Galiullina, Graduate Student, e-mail: GAF1205@yandex.ru,
I. I. Sabitov, Graduate Student, e-mail: iskra1990@gmail.com,
Ufa State Aviation Technical University (USATU), Ufa

Efficiency Evaluation of Public Services Provision Using Fuzzy Inference System

To determine the efficiency of public services provision system fuzzy inference system application based on fuzzy set theory is considered in this article. The public services provision system is considered as a queuing system. To determine the efficiency of public services provision system to citizens three of the most important indicators or parameters of queuing system are used. These parameters are service time, service quality and service cost. The model of queuing system efficiency evaluation is developed in the software package MATLAB/Simulink. This model uses a fuzzy inference system and when system efficiency is low model allows to evaluate the control action effect in a feedback on the queuing system efficiency by reducing service time and resources cost for the service.

Keywords: public services provision system, queuing system, queuing system efficiency, fuzzy inference system

References

1. **Sovetov B. Ya., Yakovlev S. A.** *Modelirovanie system* [System modeling]: ucheb. dlya vuzov. Moscow: Vysh. shkola, 2009. 344 p. (in Russian).
2. **Hinchin A. Ya.** *Raboty po matematicheskoy teorii massovogo obslugivaniya* [Works on the mathematical theory of mass service]. Moscow: Editorial URSS, 2009. 240 p. (in Russian).
3. **Chernyakhovskaya L. R.** *Modelirovanie i analiz kharakteristik dinamicheskikh processov v usloviakh neopredelennosti s primeneniem teorii massovogo obslugivaniya* [Modeling and analysis of dynamic processes characteristics under uncertainty using the mass service theory]: ucheb. posobie. Ufa: UGATU, 2014. 159 p. (in Russian).
4. **GOST R ISO 9000—2008.** *The quality management system. Basic provisions and dictionary.* National standard of the Russian Federation. Moscow: Standardinform, 2009. 35 p. (in Russian).
5. **Galiullina A. F.** Imitacionnoe modelirovanie delovykh processov kak system massovogo obslugivaniya [Simulation modeling of business processes as queuing systems], *Aktualnye problem nauki i tekhniki: Proceedings of the 9th Russian winter school-seminar of young scientists and post-graduate students.* Ufa: UGATU, 2014, vol. 1, pp. 226—230. (in Russian).
6. **Galiullina A. F.** Ocenka effektivnosti upravleniya proizvodstvennym processom s primeneniem imitacionnogo modelirovaniya na osnove system massovogo obslugivaniya [Effectiveness evaluation of production process control using simulation modeling based on queuing systems], *Vestnik UGATU.* Ufa: UGATU, 2015, vol. 19, no. 1 (67), pp. 115—122 (in Russian).
7. **Galiullina A. F.** Analiticheskoe modelirovanie delovykh processov kak system massovogo obslugivaniya dlya ocenki ikh pokazateley effektivnosti [Analytical modeling of business processes as queuing systems for assessment of their efficiency indicators], *Nauka vchera, segodnya, zavtra* [Science yesterday, today, tomorrow]: Sb. st. po materialam XX megdunar. nauch.-prakt. konf., Novosibirsk: SibAK, 2015, no. 1 (18), pp. 11—15. (in Russian).
8. **Galiullina A. F.** Modelirovanie processa predostavleniya gosudarstvennykh uslug [Modeling of the public services providing process], *Nauchnoe obozrenie*, 2015, no. 22, pp. 253—263. (In Russian).
9. **Leonenkov A. V.** Nechetkoe modelirovanie v srede Matlab i fuzzyTECH [Fuzzy modeling in Matlab and fuzzyTECH]. Saint-Petersburg: BHV — Petersburg, 2003. 736 p. (in Russian).

УДК 519.61

Ю. А. Зак, д-р техн. наук, Докт.-Ing. Deutschland, науч. консультант, e-mail: yuriy_zack@hotmail.com,
г. Аахен, Германия

Алгоритмы решения транспортной задачи в условиях нечетких данных о стоимости поставки грузов

Рассматривается математическая модель транспортной задачи линейного программирования в условиях, когда удельные транспортные расходы представлены нечеткими множествами, а объемы поставок грузов — суть действительные числа. На основе предложенных методов сравнения и определения эффективности Fuzzy-множеств общего вида разработаны нечеткие аналоги метода потенциалов. Для Fuzzy-множеств с функцией принадлежности треугольного и трапециевидного вида приведены простые формулы и подробные алгоритмы расчета, которые проиллюстрированы числовым примером.

Ключевые слова: транспортная задача, базовые допустимые решения, метод потенциалов, Fuzzy-множества, сравнение и определение эффективности нечетких множеств

Введение

Транспортная задача в традиционной постановке формулируется следующим образом [1—4]. Одно-

родный груз, находящийся на складах n поставщиков в объемах $A_1, A_2, \dots, A_j, \dots, A_n$, необходимо доставить m потребителям в объемах $B_1, B_2, \dots, B_j, \dots, B_m$. Известны транспортные расходы c_{ij} , связанные

с перевозкой единицы продукта из пункта \bar{A}_i в пункт \bar{B}_j . Рассматриваются случаи, когда

$$\sum_{i=1}^n A_i = \sum_{j=1}^m B_j.$$

Необходимо найти значение переменных $x_{ij} \geq 0$, $i = 1, \dots, m, j = 1, \dots, n$, чтобы удовлетворить спрос всех потребителей, обеспечив минимальную суммарную стоимость всех перевозок.

Математическая модель задачи может быть представлена в следующем виде:

$$\sum_{i=1}^n \sum_{j=1}^m c_{ij} x_{ij} \rightarrow \min, \quad (1)$$

где $c_{ij} \geq 0$ — некоторые заданные детерминированные действительные числа, в условиях следующей системы ограничений:

- суммарное количество продукта, направляемого i -м поставщиком всем потребителям (во все пункты назначения), должно быть равно запасу продукта в данном пункте

$$\sum_{j=1}^m x_{ij} = A_i, \quad i = 1, \dots, n; \quad (2)$$

- суммарное количество груза, доставляемого в каждый пункт назначения из всех пунктов отправления, должно быть равно его потребности

$$\sum_{i=1}^n x_{ij} = B_j, \quad j = 1, \dots, m. \quad (3)$$

Переменные задачи — это объемы перевозок $x_{ij} \geq 0$ — суть неотрицательные числа.

Данная формулировка транспортной задачи получила в литературе название замкнутой транспортной модели. Когда суммарный объем предложений (грузов, имеющихся в пунктах отправления) не равен общему объему спроса потребителей, транспортная задача называется *несбалансированной (открытой)*. Для открытой модели представляют интерес два случая:

а) суммарные запасы превышают суммарные потребности:

$$\sum_{i=1}^n A_i > \sum_{j=1}^m B_j; \quad (4)$$

б) суммарные потребности в поставке продукта превышают суммарные запасы:

$$\sum_{i=1}^n A_i < \sum_{j=1}^m B_j. \quad (5)$$

Функция цели задачи в этих обоих случаях остается неизменной. Меняется только вид системы ограничений. Открытая модель решается приведением к закрытой модели [1–4] введением фиктивного поставщика \bar{A}_{n+1} , запасы которого равны

$$A_{n+1} = \sum_{i=1}^n A_i - \sum_{j=1}^m B_j, \text{ или фиктивного потребителя } \bar{B}_{m+1}, \text{ потребность в поставках для которого равна } B_{m+1} = \sum_{j=1}^m B_j - \sum_{i=1}^n A_i.$$

При этом стоимость перевозки единицы груза как до фиктивного потребителя, так и от фиктивного поставщика полагаются равными нулю. После этих преобразований задача принимает вид закрытой модели и решается теми же алгоритмами и методами.

Решение транспортной задачи может быть получено алгоритмами симплекс-метода линейного программирования [1]. Однако ввиду практической важности этой задачи и специфики ее ограничений были разработаны более простые методы решения, позволяющие эффективно решать транспортные задачи большой размерности при существенно меньшем объеме вычислений. Одним из достаточно распространенных методов решения таких задач является метод потенциалов (см., например, [1, 3, 4]).

Однако во многих практических приложениях транспортные расходы c_{ij} , определяющие стоимость перевозки единицы груза, являются случайными величинами или могут быть представлены некоторыми нечеткими множествами. Это связано с колебанием цен на бензин и дизельное топливо, с временем года и суток, погодными условиями, типом и техническим состоянием транспортных средств и состоянием дорог, квалификацией обслуживающего персонала и многими другими причинами. В этих условиях является целесообразным суммарную стоимость перевозки грузов рассматривать и представлять в виде некоторой функции распределения или *Fuzzy*-множества, и транспортную задачу рассматривать как задачу поиска наиболее эффективной (с точки зрения минимальных затрат) функции распределения или *Fuzzy*-множества определенного вида. Для решения этой задачи могут использоваться методы *Fuzzy*-регрессионного анализа (см., например, [9–13]).

В данной работе предлагается математическая модель задачи в условиях, когда удельные транспортные расходы c_{ij} представлены нечеткими множествами, а объемы грузов на складах поставщиков $B_j, j = 1, \dots, m$, и потребности потребителей $A_i, i = 1, \dots, n$, — суть действительные числа. На основе предложенных методов сравнения эффективности и определения перспективности *Fuzzy*-множеств разработаны нечеткие аналоги метода потенциалов для *Fuzzy*-множеств с функцией принадлежности общего вида, которые детально проиллюстрированы для нечетких множеств с функцией принадлежности треугольного вида на числовом примере.

Представление суммарных затрат, связанных с доставкой грузов, в виде функции распределения

или нечеткого множества позволит более объективно оценить возможные риски, связанные с принятием решений, и обосновано принять соответствующий план поставок.

1. Постановка и математическая модель задачи

Рассматривается задача определения значения вектора действительных переменных в $x_{ij} \geq 0, i = 1, \dots, n, j = 1, \dots, m$, при условиях системы ограничений (2), (3), обеспечивающих наиболее эффективное значение (в смысле минимального значения) *Fuzzy*-множества целевой функции

$$\bar{F} = \sum_{i=1}^n \sum_{j=1}^m \bar{C}_{ij} x_{ij} \rightarrow \min, \quad (6)$$

где $\bar{C}_{ij} \geq 0$ — нечеткие множества.

В качестве *Fuzzy*-множеств можно рассматривать компактные множества с функциями принадлежности самого общего вида $0 \leq \mu_{\bar{C}}(\bar{C}_{ij}) \leq 1$. В частности, можно рассматривать функции принадлежности, представленные двумя функциями $\mu_{L(\bar{C})}(\bar{C}_{ij})$ и $\mu_{R(\bar{C})}(\bar{C}_{ij})$ или тремя функциями $\mu_{L(\bar{C})}(\bar{C}_{ij})$, $\mu_{G(\bar{C})}(\bar{C}_{ij})$ и $\mu_{R(\bar{C})}(\bar{C}_{ij})$. Функция $\mu_{L(\bar{C})}(\bar{C}_{ij})$ — монотонно неубывающая на отрезке $[0, 1]$, функция $\mu_{R(\bar{C})}(\bar{C}_{ij})$ — монотонно невозрастающая на отрезке $[0, 1]$, а функция $\mu_{G(\bar{C})}(\bar{C}_{ij}) = \text{const} = 1$. Пример таких функций представлен на рис. 1.

Так как значение целевой функции в допустимых (базисных опорных) и оптимальном решениях также будет представлено нечетким множеством, рассмотрим методы сравнения и определения пред-

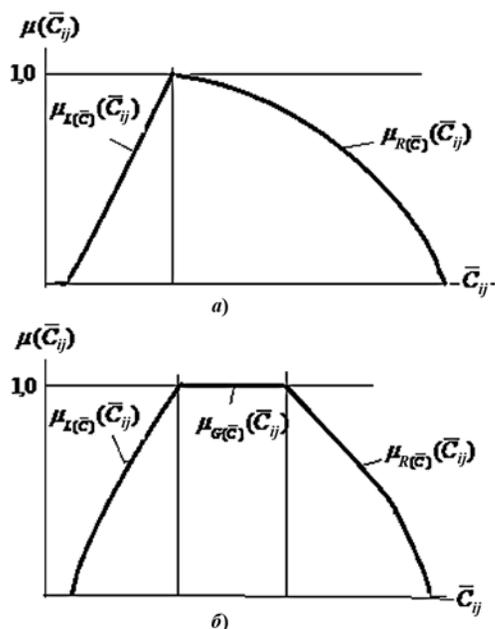


Рис. 1

почтений *Fuzzy*-множеств. Алгоритмы сравнения, определения абсолютных и относительных предпочтений и ранжирования *Fuzzy*-множеств были рассмотрены во многих публикациях отечественной и зарубежной литературы (см., например [7, 8]), а также в статьях и монографии автора для функций принадлежности специального вида [5, 6].

Ниже приведены алгоритмы сравнения и определения предпочтений нечетких множеств \bar{C}_1 и \bar{C}_2 с функциями принадлежности $\mu_{\bar{C}}(\bar{C}_1)$ и $\mu_{\bar{C}}(\bar{C}_2)$ описанного выше самого общего вида, которые использованы в алгоритмах решения сформулированной задачи. Обозначим $\bar{C}_1 <_{\text{Re}} \bar{C}_2$ факт, что *Fuzzy*-множество \bar{C}_1 является более предпочтительным (в смысле наименьшего значения), чем *Fuzzy*-множество \bar{C}_2 . Предложенные ниже критерии и алгоритмы сравнения позволяют на основе вычисления некоторого детерминированного эквивалента нечеткого множества определить правила относительного предпочтения.

2. Алгоритмы сравнения и определения предпочтений нечетких множеств

Метод, основанный на анализе сечений нечетких множеств. Этот метод, который достаточно подробно описан во многих статьях и монографиях, проиллюстрирован рис. 2 и заключается в следующем.

Рассматриваются значения некоторого числа сечений нечетких множеств $(\bar{C}, \mu_{\bar{C}}(C)) - \alpha_0, \alpha_1, \dots, \alpha_p, \dots, \alpha_p$. Пусть $C_0^1, C_1^1, \dots, C_p^1, \dots, C_p^1$ и $C_0^2, C_1^2, \dots, C_p^2, \dots, C_p^2$ — соответственно координаты левых и правых крайних точек координаты абсцисс всех этих сечений. Вычислим значения координат средних точек соответствующих сечений

$$L_p = 0,5 \cdot (C_p^1 + C_p^2). \quad (7)$$

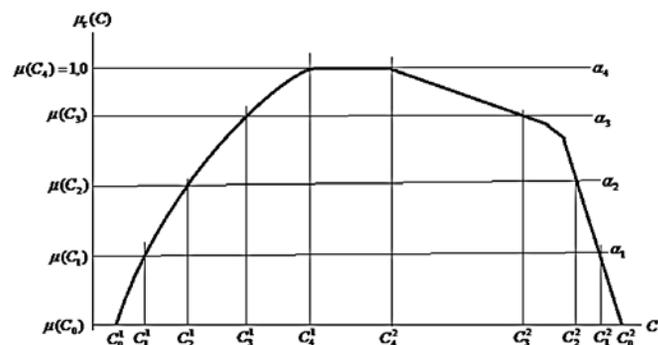


Рис. 2

Определив значения весовых коэффициентов из условий

$$0 \leq w_p \leq 1, p = 0, 1, \dots, P; \sum_{p=0}^P w_p = 1, \quad (8)$$

вычислим некоторое средневзвешенное детерминированное значение эффективности данного Fuzzy-множества:

$$S(\bar{C}) = \sum_{p=0}^P w_p L_p. \quad (9)$$

На основании вычисленных для двух Fuzzy-множеств $(\bar{C}_1, \mu_{\bar{C}}(C_1))$ и $(\bar{C}_2, \mu_{\bar{C}}(C_2))$ значений $S(\bar{C}_1)$ и $S(\bar{C}_2)$ можно заключить

$$(\bar{C}_1, \mu_{\bar{C}}(C_1)) <_{\text{Re}} (\bar{C}_2, \mu_{\bar{C}}(C_2)), \quad \text{если } S(\bar{C}_1) \leq S(\bar{C}_2). \quad (10)$$

Отметим, что выбор числа и уровня сечений, а также значений весовых коэффициентов связан с предпочтениями экспертов и лица, принимающего решение, степенью допускаемого риска. Поэтому определение эффективности Fuzzy-множеств этим методом является в некоторой степени субъективным. Альтернативным методом сравнения и определения предпочтений является также метод взвешенной точки, который учитывает все α -срезы функций принадлежности [9, 10] и может также быть использован в приведенных ниже алгоритмах.

Определение предпочтений сравнением координаты абсциссы центра тяжести нечетких множеств. Следует отметить, что выбор числа сечений (значения P) уровня сечений (значений $\mu_{\bar{C}}(C_p)$, $p = 0, 1, \dots, P$), значений весовых коэффициентов w_p является субъективным и определяется лицом, принимающим решение. Более объективным методом определения относительных предпочтений нечетких множеств является алгоритм, основанный на сравнении абсциссы их центров тяжести, которые вычисляют по формуле

$$G(\bar{C}) = \frac{\int_{-\infty}^{\infty} C \mu_{\bar{C}}(C) dC}{\int_{-\infty}^{\infty} \mu_{\bar{C}}(C) dC}. \quad (11)$$

Следовательно,

$$(\bar{C}_1, \mu_{\bar{C}}(C_1)) \leq_{\text{Re}} (\bar{C}_2, \mu_{\bar{C}}(C_2)), \quad \text{если } G(\bar{C}_1) \leq G(\bar{C}_2). \quad (12)$$

Сравнение нечетких множеств трапециевидного и треугольного вида. Рассмотрим нечеткие множества трапециевидного и треугольного видов

$\bar{C} = (a(\bar{C}), m_1(\bar{C}), m_2(\bar{C}), b(\bar{C}))$ и $\bar{C} = (a(\bar{C}), m(\bar{C}), b(\bar{C}))$ (см. рис. 2).

Наиболее объективным методом определения предпочтений нечетких множеств в смысле $(\bar{C}_1, \mu_{\bar{C}}(C_1)) \leq_{\text{Re}} (\bar{C}_2, \mu_{\bar{C}}(C_2))$ является проверка выполнения системы неравенств:

- для функций принадлежности трапециевидного вида

$$\begin{aligned} (m_1(C_1) - a(C_1)) &\leq (m_1(C_2) - a(C_2)), \\ m_1(C_1) &\leq m_1(C_2), m_2(C_1) \leq m_2(C_2), \\ (m_2(C_1) + b(C_1)) &\leq (m_2(C_2) + b(C_2)); \end{aligned} \quad (13)$$

- для функций принадлежности треугольного вида
- $$\begin{aligned} (m(C_1) - a(C_1)) &\leq (m(C_2) - a(C_2)), m(C_1) \leq m(C_2), \\ (m(C_1) + b(C_1)) &\leq (m(C_2) + b(C_2)), \end{aligned} \quad (14)$$

и хотя бы одно из этих неравенств выполняется как строгое неравенство.

Значения $S(\bar{C})$ и $G(\bar{C})$ для частных случаев функций принадлежности нечетких множеств трапециевидного и треугольного видов (рис. 3) вычисляются по формулам (9), (11), а значения L_p определяются соответственно выражениями

$$L_p = 0,5\{(m_1(\bar{C}) - a(\bar{C})(1 - \alpha_p)) + (m_2(\bar{C}) + b(\bar{C})(1 - \alpha_p)) + (m_1(\bar{C}) + m_2(\bar{C}))\}; \quad (15)$$

$$L_p = 0,5\{(m(\bar{C}) - a(\bar{C})(1 - \alpha_p)) + (m(\bar{C}) + b(\bar{C})(1 - \alpha_p))\}. \quad (16)$$

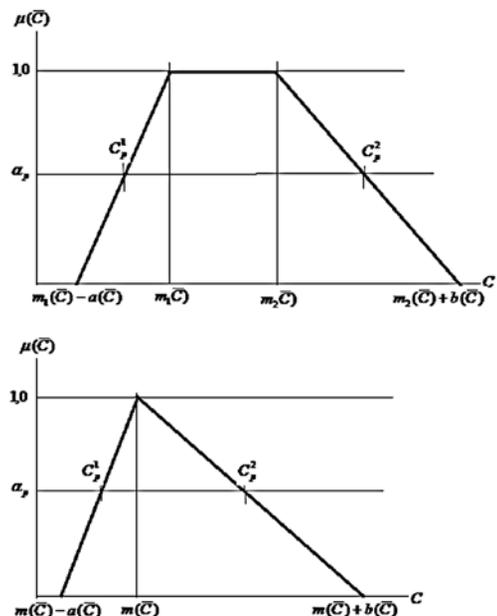


Рис. 3

Аналоги значения $G(\bar{C})$ в этом случае определяются по следующим явным формулам:

- для функций принадлежности треугольного вида

$$G(\bar{C}) = \frac{1}{3} [(m(\bar{C}) - a(\bar{C})) + m(\bar{C}) + (b(\bar{C}) + m(\bar{C}))]; \quad (17)$$

- для функций принадлежности трапециевидного вида может рассматриваться показатель

$$G_1(\bar{C}) = \frac{1}{3} [(m_1(\bar{C}) - a(\bar{C})) + m_1(\bar{C}) + (b(\bar{C}) + m_1(\bar{C}))] + \frac{1}{2} (m_1(\bar{C}) + m_2(\bar{C})). \quad (18)$$

3. Алгоритмы получения базового опорного решения

Как и для любой задачи линейного программирования, оптимальный план транспортной задачи является также ее опорным планом. Опорные и оптимальные планы решения задачи сводятся в таблицу, в каждой клетке которой определен объем поставок из пункта \bar{A}_i в пункт \bar{B}_j , т.е. значение переменной x_{ij} .

Решение транспортной задачи состоит из двух этапов. 1-й этап предусматривает нахождение ее какого-либо опорного плана. Определение опорного плана может быть выполнено одним из следующих трех методов [1, 3]:

- метод северо-западного угла;
- метод минимального элемента;
- метод аппроксимации Фогеля.

Рассмотрим первые два из этих методов.

Алгоритм получения опорного решения методом северо-западного угла ничем не отличается от применяемого алгоритма в случае детерминированных данных коэффициентов целевой функции (6).

Получение опорного решения методом минимального элемента предусматривает выполнение следующих процедур:

- расчет детерминированного эквивалента для каждого нечеткого множества коэффициента целевой функции \bar{C}_{ij} в виде $S(\bar{C}_{ij})$ по формулам (7)–(9), для функций принадлежности трапециевидного и треугольного вида проверкой системы неравенств (13), (14) либо вычисления значения L_p по явным формулам (15), (16), либо в виде значения центра тяжести соответствующего *Fuzzy*-множества по формуле (17), (18);
- так как вычисленные оценки являются детерминированными величинами, упорядочим их в соответствии с выражениями (10) или (12) по возрастанию вычисленных значений, после чего все шаги получения опорного решения ничем не отличаются от последовательности вычислений в случае детерминированных значений c_{ij} .

Пусть $\bar{x}_{ij} \geq 0$ — детерминированные значения объемов перевозок, полученные в некотором до-

пустимом базовом решении. *Fuzzy*-множество, определяющее значение целевой функции полученного базисного решения, определяется по формуле

$$\begin{aligned} & \mu_{\bar{F}(\bar{Z})}(\bar{F}(\bar{Z}_p)) = \\ & = \max_{\bar{x}_{11}, \bar{x}_{12}, \dots, \bar{x}_{ij}, \dots, \bar{x}_{nm}} \min_{\bar{x}_{ij} \in \tilde{Q}(\bar{Z}_p)} \{ \mu_{\bar{C}}(\bar{C}_{11}), \mu_{\bar{C}}(\bar{C}_{12}), \dots, \\ & \mu_{\bar{C}}(\bar{C}_{ij}), \dots, \mu_{\bar{C}}(\bar{C}_{nm}) \}, \end{aligned} \quad (19)$$

где

$$\begin{aligned} & \tilde{Q}(\bar{Z}_p) = \\ & = \left\{ (\bar{x}_{11}, \bar{x}_{12}, \dots, \bar{x}_{ij}, \dots, \bar{x}_{nm}) \mid \sum_{i=1}^n \sum_{j=1}^m \bar{C}_{ij} \bar{x}_{ij} = \bar{Z}_p; \right. \\ & \left. \bar{x}_{ij} > 0, i = 1, \dots, n, j = 1, \dots, m \right\}; \end{aligned} \quad (20)$$

$\bar{Z}_p \in [\bar{Z}^{\min}, \bar{Z}^{\max}]$ — диапазон возможных детерминированных значений суммарных затрат, связанных с перевозкой грузов.

Для функций принадлежности удельных затрат доставки грузов трапециевидного или треугольного вида выражение $\mu_{\bar{F}(\bar{Z})}(\bar{F}(\bar{Z}))$ вычисляется по явным формулам

$$\begin{aligned} & \mu_{\bar{F}(\bar{Z})}(\bar{F}(\bar{Z})) \rightarrow (\bar{A}, \bar{M}_1, \bar{M}_2, \bar{B}) \\ & \text{или } \mu_{\bar{F}(\bar{Z})}(\bar{F}(\bar{Z})) \rightarrow (\bar{A}, \bar{M}, \bar{B}), \end{aligned} \quad (21)$$

где

$$\begin{aligned} & \bar{A} = \left\{ \sum_{i=1}^n \sum_{j=1}^m a(\bar{C}_{ij}) \bar{x}_{ij} \mid \bar{x}_{ij} > 0, \right. \\ & \left. i = 1, \dots, n, j = 1, \dots, m \right\}; \end{aligned} \quad (22)$$

$$\begin{aligned} & \bar{B} = \left\{ \sum_{i=1}^n \sum_{j=1}^m b(\bar{C}_{ij}) \bar{x}_{ij} \mid \bar{x}_{ij} > 0, \right. \\ & \left. i = 1, \dots, n, j = 1, \dots, m \right\}; \end{aligned} \quad (23)$$

$$\begin{aligned} & \bar{M}_k = \left\{ \sum_{i=1}^n \sum_{j=1}^m m_k(\bar{C}_{ij}) \bar{x}_{ij} \mid \bar{x}_{ij} > 0, \right. \\ & \left. i = 1, \dots, n, j = 1, \dots, m \right\}, k = 1, 2; \end{aligned} \quad (24)$$

$$\begin{aligned} & \bar{M} = \left\{ \sum_{i=1}^n \sum_{j=1}^m m(\bar{C}_{ij}) \bar{x}_{ij} \mid \bar{x}_{ij} > 0, \right. \\ & \left. i = 1, \dots, n, j = 1, \dots, m \right\}. \end{aligned} \quad (25)$$

4. Улучшение опорного плана методом потенциалов

Построенный план перевозок является допустимым, но не является оптимальным. Для нахождения оптимального плана перевозок необходимо последовательно, пока это возможно, улучшать опорный план путем отыскания на каждом шаге выгодных циклов переноса грузов. Цикл позволяет перераспределить занятые клетки так, чтобы получить новый план перевозок с меньшими суммарными затратами. На каждом шаге (цикле) алгоритма в условиях детерминированных значений стоимостей перевозок c_{ij} заменяют одну свободную переменную на базисную, т. е. заполняют одну клетку и взамен этого освобождают одну из других базисных клеток. Ниже рассматривается реализация этих процедур в условиях нечетких данных о стоимости перевозок, заданными *Fuzzy*-множествами \bar{C}_{ij} с функциями принадлежности $0 \leq \mu_{\bar{C}}(\bar{C}_{ij}) \leq 1$.

Расчет потенциалов

$$\mu_{\bar{U}}(\bar{U}_i^r) = \max_{(\bar{C}_{ij}, \bar{V}_j) \in \tilde{H}_1(\bar{U}_i^r)} \{\mu_{\bar{C}}(\bar{C}_{ij}), \mu_{\bar{V}}(\bar{V}_j^l)\}, \quad (26)$$

где

$$\begin{aligned} & \tilde{H}_1(\bar{U}_i^r) = \\ & = \{C_{ij}^t, t=1, \dots, T; \bar{V}_j^l, l=1, \dots, L \mid \bar{U}_i^r = \bar{C}_{ij}^t - \bar{V}_j^l\}; \end{aligned} \quad (27)$$

$$\mu_{\bar{V}}(\bar{V}_j^l) = \max_{(\bar{C}_{ij}, \bar{U}_i) \in \tilde{H}_2(\bar{V}_j^l)} \{\mu_{\bar{C}}(\bar{C}_{ij}), \mu_{\bar{U}}(\bar{U}_i)\}; \quad (28)$$

$$\begin{aligned} \tilde{H}_2(\bar{V}_j^l) & = \{C_{ij}^t, t=1, \dots, T; \bar{U}_i^r, r=1, \dots, R \mid \bar{V}_j^l = \\ & = \bar{C}_{ij}^t - \bar{U}_i^r\}. \end{aligned} \quad (29)$$

Здесь $\bar{U}_i^r \in [\bar{U}_i^{\min}, \bar{U}_i^{\max}]$, $\bar{V}_j^l \in [\bar{V}_j^{\min}, \bar{V}_j^{\max}]$ — соответственно диапазоны возможных детерминированных значений соответствующих потенциалов строк и столбцов.

Значения свободных (соответствующих значениям $x_{ij} = 0$) клеток таблицы вычисляют по формулам

$$\begin{aligned} & \mu_{\bar{W}}(\bar{W}_{ij}^k) = \\ & = \max_{(\bar{C}_{ij}, \bar{U}_i, \bar{V}_j) \in \tilde{\Omega}(\bar{W}_{ij}^k)} \{\mu_{\bar{C}}(\bar{C}_{ij}), \mu_{\bar{U}}(\bar{U}_i^r), \mu_{\bar{V}}(\bar{V}_j^l) \mid x_{ij} = \\ & \quad 0\}, \end{aligned} \quad (30)$$

где

$$\begin{aligned} \tilde{\Omega}(\bar{W}_{ij}^k) & = \{C_{ij}^t, t=1, \dots, T; \bar{U}_i^r, r=1, \dots, R, \\ & \bar{V}_j^l, l=1, \dots, L \mid \bar{W}_{ij}^k = \bar{C}_{ij}^t - \bar{U}_i^r - \bar{V}_j^l\}. \end{aligned} \quad (31)$$

Здесь $\bar{W}_{ij}^k \in [\bar{W}_{ij}^{\min}, \bar{W}_{ij}^{\max}]$ — диапазоны возможных детерминированных значений соответствующих потенциалов свободных клеток таблицы.

В случае функций принадлежности трапециевидного и треугольного видов

$$\begin{aligned} \bar{C} & = (a(\bar{C}), m_1(\bar{C}), m_2(\bar{C}), b(\bar{C})) \\ & \text{и } \bar{C} = (a(\bar{C}), m(\bar{C}), b(\bar{C})) \end{aligned}$$

вычисления потенциалов строк и столбцов и оценок свободных клеток выполняют по формулам

$$\begin{aligned} \bar{U}_i & = (a(\bar{C}_{ij}) + b(\bar{V}_j), m_1(\bar{C}) - m_1(\bar{V}_j), m_2(\bar{C}) - \\ & \quad - m_1(\bar{V}_j), b(\bar{C}) + a(\bar{V}_j)), \end{aligned} \quad (32)$$

$$\bar{U}_i = (a(\bar{C}_{ij}) + b(\bar{V}_j), m(\bar{C}) - m(\bar{V}_j), b(\bar{C}) + a(\bar{V}_j)); \quad (33)$$

$$\begin{aligned} \bar{W}_{ij} & = (a(\bar{C}_{ij}) + a(\bar{U}_i) + b(\bar{V}_j), m_1(\bar{C}) - \\ & \quad - m_1(\bar{U}_i) - m_1(\bar{V}_j), m_2(\bar{C}) - m_1(\bar{U}_i) - \\ & \quad - m_1(\bar{V}_j), b(\bar{C}) + a(\bar{U}_i) + a(\bar{V}_j)), \end{aligned} \quad (34)$$

$$\begin{aligned} \bar{W}_{ij} & = (a(\bar{C}_{ij}) + a(\bar{U}_i) + b(\bar{V}_j), m(\bar{C}) - m(\bar{U}_i) - \\ & \quad - m(\bar{V}_j), -m_1(\bar{V}_j), b(\bar{C}) + a(\bar{U}_i) + a(\bar{V}_j)). \end{aligned} \quad (35)$$

На основе вычисленных по приведенным выше формулам (7)–(9) функций принадлежности оценок свободных клеток таблицы (а в случае трапециевидных или треугольных функций принадлежности по формулам (13), (14)) вычисляют детерминированные оценки эффективности нечетких множеств $S(\bar{C}_{ij})$, либо координаты оси абсцисс их центров тяжести $G(\bar{C}_{ij})$ в соответствии с выражением (11) (или (17), (18)). На основе вычисленных детерминированных оценок дальнейшие шаги преобразования и улучшения опорного плана выполняются как и в методе потенциалов в условиях детерминированных данных об удельных стоимостях перевозок.

Отметим, что в приближенных методах решения задачи значения потенциалов и оценок свободных клеток можно определять по более простым формулам на основе детерминированных оценок эффективности *Fuzzy*-множеств:

$$\begin{aligned} \bar{U}_i & = S(\bar{C}_{ij}) - S(\bar{V}_j), \quad \bar{V}_j = S(\bar{C}_{ij}) - S(\bar{U}_i), \\ \bar{W}_{ij} & = S(\bar{C}_{ij}) - S(\bar{U}_i) - S(\bar{V}_j) \end{aligned} \quad (36)$$

или

$$\begin{aligned} \bar{U}_i & = G(\bar{C}_{ij}) - G(\bar{V}_j), \quad \bar{V}_j = G(\bar{C}_{ij}) - G(\bar{U}_i), \\ \bar{W}_{ij} & = G(\bar{C}_{ij}) - G(\bar{U}_i) - G(\bar{V}_j). \end{aligned} \quad (37)$$

5. Иллюстративный пример

Требуемые объемы поставок потребителям и запасы товаров поставщиков заданы соответственно в последней строке и последнем столбце табл. 1. Стоимости перевозки единицы объема грузов заданы *Fuzzy*-числами с функциями принадлежности треугольного вида $\mu_X(A) = (a, m, b)$ и представлены в клетках табл. 1. Необходимо найти наиболее эффективный план перевозок, обеспечивающий наи-

лучшее значение *Fuzzy*-множества, определяющего суммарные затраты на выполнение плана перевозок.

Сравнение и определение предпочтений *Fuzzy*-множеств выполняется либо координатой абсцисс центра тяжести, либо методом сечений на уровнях $\lambda_0\{\mu_X(A)\} = 0, \lambda_1\{\mu_X(A)\} = 0,25, \lambda_2\{\mu_X(A)\} = 0,5, \lambda_3\{\mu_X(A)\} = 0,8, \lambda_4\{\mu_X(A)\} = 1,0$, а соответствующие весовые коэффициенты при расчете детерминированного показателя эффективности *Fuzzy*-множеств заданы вектором $W = (0,05; 0,1; 0,15; 0,3; 0,4)$.

Таблица 1

Исходные данные задачи

	1	2	3	4	A_i
1	(1, 4, 2)	(2, 3, 1)	(2, 6, 3)	(1, 2, 3)	10
2	(2, 5, 1)	(1, 6, 3)	(1, 7, 2)	(1, 4, 3)	15
3	(2, 4, 3)	(2, 3, 3)	(1, 5, 2)	(1, 6, 2)	12
B_j	15	8	6	8	37

Таблица 2

Получение опорного решения методом северо-западного угла (детерминированные оценки стоимости получены методом сечений)

	1	2	3	4	A_i
1	4,13 10	2,87	6,13	2,23	10
2	4,87	6,11 8	7,13	4,26	15
3	4,13	3,13	5,13	6,13 8	12
B_j	15	8	6	8	37

Таблица 3

Получение опорного решения методом выбора минимального элемента (детерминированные оценки стоимости получены методом сечений)

	1	2	3	4	A_i
1	4,13	2,87 2	6,13	2,23 8	10
2	4,87	6,11	7,13	4,26 6	15
3	4,13 6	3,13 6	5,13	6,13	12
B_j	15	8	6	8	37

Таблица 4

Получение опорного решения методом выбора минимального элемента (детерминированные оценки стоимости как центр тяжести соответствующих подмножеств)

	1	2	3	4	A_i
1	4,333	2,667 8	6,333	2,667 2	10
2	4,667	6,667	7,333	4,667 6	15
3	4,333 12	3,333	5,333	6,333	12
B_j	15	8	6	8	37

Fuzzy-множество значения целевой функции опорного плана (табл. 2)

$$10*(1, 4, 2) + 5*(2, 5, 1) + 8*(1, 6, 3) + 2*(1, 7, 2) + 4*(1, 5, 2) + 8*(1, 6, 2) = (42, 195, 77),$$

т. е. $a = 42; m = 195; b = 77$.

$$F_0^0 = (m(\bar{C}) - a(\bar{C}), m(\bar{C}), m(\bar{C}) + b(\bar{C})) = (153, 195, 272),$$

$$C(F_0^0) = 0,05 \cdot \frac{1}{2} \cdot [(195 - 42 \cdot (1,0 - 0)) + (195 + 77 \cdot (1,0 - 0))] + 0,1 \cdot \frac{1}{2} \cdot [(195 - 42 \cdot (1,0 - 0,25)) + (195 + 77 \cdot (1,0 - 0,25))] + 0,15 \cdot \frac{1}{2} \cdot [(195 - 42 \cdot (1,0 - 0,5)) + (195 + 77 \cdot (1,0 - 0,5))] + 0,3 \cdot \frac{1}{2} \cdot [(195 - 42 \cdot (1,0 - 0,8)) + (195 + 77 \cdot (1,0 - 0,8))] + 0,4 \cdot \frac{1}{2} \cdot [(195 - 42 \cdot (1,0 - 1,0)) + (195 + 77 \cdot (1,0 - 1,0))] = 199,49.$$

$$G(F_0^0) = \frac{1}{3} \cdot (153 + 195 + 272) = 206,67.$$

Значения параметров F_0^0 показаны на рис. 3. Аналогичным образом вычислены параметры *Fuzzy*-множеств F_0^1, F_0^2 и F_0^3 .

Fuzzy-множество значения целевой функции опорного плана (табл. 3):

$$2*(2, 3, 1) + 8*(1, 2, 3) + 9*(2, 5, 1) + 6*(1, 7, 2) + 6*(2, 4, 3) + 6*(2, 3, 3) = (60, 151, 83).$$

$$F_0^1 = (91, 151, 234);$$

$$C(F_0^1) = 153,985; G(F_0^1) = 158,67.$$

Так как $91 < 153, 151 < 195, 234 < 272$, то $F_0^1 <_{RE} F_0^0$. *Fuzzy*-множество значения целевой функции опорного плана (табл. 4):

$$8*(2, 3, 1) + 2*(1, 2, 3) + 3*(2, 5, 1) + 6*(1, 7, 2) + 6*(1, 4, 3) + 12*(2, 4, 3) = (60, 157, 101);$$

$$F_0^2 = (97, 157, 258); C(F_0^2) = 162,32; G(F_0^2) = 170,67.$$

Так как $97 < 153, 157 < 195, 258 < 272$, то $F_0^2 <_{RE} F_0^0$.

Как и по условиям $C(F_0^1) < C(F_0^2)$, так и при сравнении *Fuzzy*-множеств по критерию координаты абсцисс центра тяжести установлено, что $F_0^1 <_{RE} F_0^2$.

Рассмотрим наилучший из допустимых опорных планов, представленный в табл. 3. Рассчитаем значения потенциалов и оценок свободных клеток табл. 3 и результаты расчетов представим в табл. 5.

$$u_1 = 0, v_2 = c_{12} - u_1 = 2,87, v_4 = c_{14} - u_1 = 2,26, u_3 = c_{32} - v_2 = 0,26, v_1 = c_{31} - u_3 = 3,87, u_2 = c_{23} - v_3 = 1,0, v_3 = c_{23} - u_2 = 6,13.$$

Таблица 5

Допустимое базовое решение задачи

	1	2	3	4	u_i
1	0,26	2,87	0	2,26	0
2	4,87	2,24	6,66	1,0	1,0
3	4,13	3,13	-1,26	3,64	0,26
v_j	3,87	2,87	6,13	2,26	

Таблица 6

Наиболее эффективный план перевозок

	1	2	3	4	A_i
1		2		8	10
2	15				15
3		6	6		12
B_j	15	8	6	8	

$$\tilde{c}_{11} = c_{11} - u_1 - v_1 = 4,13 - 0 - 3,87 = 0,26,$$

$$\tilde{c}_{13} = c_{13} - u_1 - v_3 = 6,13 - 0 - 6,13 = 0,$$

$$\tilde{c}_{22} = c_{22} - u_2 - v_2 = 6,11 - 1,0 - 2,87 = 2,24,$$

$$\tilde{c}_{24} = c_{24} - u_2 - v_4 = 4,26 - 1,0 - 2,26 = 1,0,$$

$$\tilde{c}_{33} = c_{33} - u_3 - v_3 = 5,13 - 0,26 - 6,13 = -1,26,$$

$$\tilde{c}_{34} = c_{34} - u_3 - v_4 = 6,13 - 0,26 - 2,2 = 3,64.$$

Так как значение элемента $\tilde{c}_{33} = -1,26$, т. е. отрицательно, то оптимальное решение задачи не получено. Строим цикл, включающий следующие клетки таблицы $(3,3)_+ \rightarrow (3,1)_- \rightarrow (2,1)_+ \rightarrow (2,3)_- \rightarrow (3,3)_+$. Минимальный элемент величины поставок клеток этого цикла равен $\lambda = \min x_{ij}$ и достигается в клетке $x_{31} = 6$. Выводим эту клетку из базисного решения, выполняем перемещение минимальной величины поставки по циклу, и пересчитываем таблицу. Полученное новое базисное решение, которое является наиболее эффективным решением данной задачи, приведено в табл. 6.

$$2*(2, 3, 1) + 8*(1, 2, 3) + 15*(2, 5, 1) + 6*(2, 3, 3) + 6*(1, 5, 2) = (60, 145, 71).$$

$$F_0^3 = (85, 145, 216); C(F_0^3) = 139,392; G(F_0^3) = 148,667.$$

Так как $85 < 91$, $145 < 151$, $216 < 234$, то $F_0^4 <_{RE} F_0^1$ и, следовательно, является самым эффективным из всех рассмотренных выше планов перевозок.

Заключение

Во многих практических приложениях в связи с колебанием цен на бензин и дизельное топливо, погодными условиями, типом и состоянием дорог

и видом используемых транспортных средств, квалификацией менеджмента и обслуживающего персонала и многими другими причинами транспортные расходы c_{ij} , определяющие стоимость перевозки единицы груза, меняются в процессе реализации плана перевозок и могут быть представлены некоторыми нечеткими множествами. Суммарную стоимость перевозки грузов в этих условиях предлагается рассматривать и представлять в виде некоторого нечеткого множества, и транспортную задачу решать как задачу поиска наиболее эффективного (с точки зрения минимальных затрат) Fuzzy-множества определенного вида.

На основе предложенных методов сравнения и определения эффективности Fuzzy-множеств общего вида разработаны нечеткие аналоги метода потенциалов. Для Fuzzy-множеств с функцией принадлежности треугольного и трапециевидного вида приведены простые формулы и подробные алгоритмы расчета, которые проиллюстрированы на числовом примере.

Представление суммарных затрат, связанных с реализацией наиболее эффективного плана перевозок, в виде нечеткого множества позволит более объективно оценить возможные риски, связанные с принятием решений, и обосновано принять полученный в результате анализа план поставок

Список литературы

1. Гольштейн Е. Г., Юдин Д. Б. Задачи линейного программирования транспортного вида. М.: Физматгиз, Наука, 1993. 384 с.
2. Канторович Л. В. О перемещении масс // Докл. АН СССР. 1942. Т. 37, № 3. С. 227–229.
3. Лунгу К. Н. Линейное программирование. Руководство к решению задач. М.: Физматлит, 2005. 128 с.
4. Данциг Дж. Линейное программирование, его применения и обобщения. М.: Прогресс, 1966. 600 с.
5. Зак Ю. А. Принятие решений в условиях размытых и нечетких данных. Fuzzy-технологии. М.: URSS, Либроком, 2013. 352 с.
6. Зак Ю. А. Критерии и методы сравнения нечетких множеств // Системные исследования и информационные технологии. 2013. № 3. С. 58–68.
7. Rommelfanger H.-J. Entscheiden bei Unschärfe. Fuzzy Decision. Support-Systeme. Berlin-Heidelberg: Springer Verlag, 1994. 314 p.
8. Duboi D., Prade H. M. Fuzzy sets and systems: theory and applications. New York — London — Toronto: Academic Press, 1980. 393 p.
9. Chang Y.-H. O. Hybrid fuzzy least-squares regression analysis and its reliability measures // Fuzzy Sets and Systems. 2001. Vol. 119. P. 225–246.
10. Poleshuk O. M., Komarov E. G. New defuzzification method based on weighted intervals // Proceedings of the 27th International Conference of the North American Fuzzy Information Processing Society, NAFIPS'2008.
11. Poleshchuk O., Komarov E. A Fuzzy Nonlinear Regression Model for Interval Type-2 Fuzzy Sets // International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering. 2014. Vol. 8, N. 6. P. 833–837.
12. Tanaka H., Uejima S., Asai K. Linear regression analysis with fuzzy model // IEEE Transactions on Systems, Man and Cybernetics. 1982. Vol. 12, N. 6. P. 903–907.
13. Зак Ю. А. Fuzzy-регрессионные модели прогнозирования затрат времени и стоимости грузовых автомобильных перевозок // Логистика сегодня. 2015. № 3. С. 162–172.

Algorithms for Solving the Transport Problem in Fuzzy Data on the Cost of Delivery of Goods

We consider a mathematical model of the transportation problem of linear programming in an environment where the specific transport costs are represented by fuzzy sets, and volume of cargo shipments — are real numbers. On the basis of the proposed methods of comparison and benchmarking fuzzy-sets general type developed fuzzy analogues of potential method. For Fuzzy-sets with the membership function of the triangular and trapezoidal species we present simple formulas and detailed calculation algorithms, which are illustrated in the numerical example.

Keywords: transportation problem, the basic feasible solutions, the potential method, Fuzzy-set comparison and determine the effectiveness of fuzzy sets

References

1. Golshtein E. G., Yudin D. B. *Zadachi linejnogo programmirovaniya transportnogo vida*, Moscow, Fismatgiz, Nauka, 1993, 384 p. (in Russian).
2. Kantorovitch L. V. O peremeschenii mass. *Doklady AN SSSR*, 1942, vol. 37. P. 227—229. (in Russian).
3. Lungu K. N. *Linejnoje programmirovaniye. Rukovodstvo k resheniyu zadach*. Moscow: Fismatlit, 2005. 128 p. (in Russian).
4. Danzing Dzh. *Linejnoje programmirovaniye, jego primeneniya i obobscheniya*, Moscow: Progress, 1966, 600 p. (in Russian).
5. Zack Yu. A. *Prinyatiye reshenij v usloviyach razmitich i nechetkikh dannich. Fuzzy-technologii*, Moscow, Librokom, 2013, 352 p.
6. Zack Yu. A. *Kriterii i metodi sravneniya nechetkikh mnozhestv, Sistemnije issledovaniya i informazionnije tehnologii*, Kiev, 2013, no. 3. pp. 58—68.
7. Rommelfanger H.-J. *Entscheiden bei Unschurfe. Fuzzy Decision. Support-Systeme*, Springer Verlag, Berlin-Heidelberg, 1994, 314 p.
8. Duboi D., Prade H. M. *Fuzzy sets and systems: theory and applications*, Academic Press, New York — London — Toronto, 1980, 393 p.
9. Chang Y.-H. O. Hybrid fuzzy least-squares regression analysis and its reliability measures, *Fuzzy Sets and Systems*, 2001, vol. 119, pp. 225—246.
10. Poleshuk O. M., Komarov E. G. New defuzzification method based on weighted intervals, *Proceedings of the 27th International Conference of the North American Fuzzy Information Processing Society, NAFIPS'2008*.
11. Poleshchuk O., Komarov E. A Fuzzy Nonlinear Regression Model for Interval Type-2 Fuzzy Sets, *International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering*, 2014, vol. 8, no. 6, pp. 833—837.
12. Tanaka H., Uejima S., Asai K. Linear regression analysis with fuzzy model, *IEEE Transactions on Systems, Man and Cybernetics*, 1982, vol. 12, no. 6, pp. 903—907.
13. Zack Yu. A. Fuzzy-regressionnije modeli prognozirovaniya zatraat vremeni i stoimosti grusovich avtomobilnich perevosok, *Logistika segodnya*, 2015, no. 3, pp. 162—172.

УДК 004.91

С. В. Бутаков, канд. техн. наук, доц., e-mail: sergey.butakov@computer.org,

С. В. Мурзинцев, аспирант, e-mail: o.100@yandex.ru,

А. А. Цхай, д-р техн. наук, проф., e-mail: taa1956@mail.ru,

Алтайская академия экономики и права, Барнаул

Использование горизонтально масштабируемой инфраструктуры при поиске заимствований в тексте

Рассмотрена проблема быстрого сравнения текстовых документов. В прикладных задачах акцент сделан на поиск плагиата и на фильтрацию текстов в системах защиты от утечек информации. Краткий обзор решений, основанных на традиционных СУБД, показал их ограничения с точки зрения масштабируемости системы. В качестве альтернативы предложено использовать нереляционные СУБД с возможностью распределения поискового индекса между узлами системы. Для решения задач текстового поиска в работе предложен вариант представления отпечатков текстов в виде "ключ — значение", выполнена программная реализация данной модели и проведены эксперименты, подтвердившие приемлемость модели с точки зрения реализации на горизонтально масштабируемой платформе.

Ключевые слова: фильтрация текста, сравнение текстов, большие данные, обнаружение плагиата, системы предотвращения потери данных

Введение

Задача поиска совпадающих или близких к совпадению фрагментов текстов изучается на протяжении нескольких последних десятилетий. В боль-

шом потоке появившихся исследований можно выделить два основных направления. В работах первого направления совпадение или близость к совпадению фрагментов текстов понимается как совпадение или близость смыслов фрагментов [1, 2].

Решение этой задачи в различных вариантах привело к созданию семейства систем семантического поиска [3, 4] и, как следствие, к улучшению поиска информации в сети Интернет.

Работы, относящиеся ко второму направлению, ориентированы на выявление совпадения или близости фрагментов текста на поверхностном, т. е. символьном, уровне. Такая постановка задачи актуальна, в частности, при поиске плагиата в текстах [5, 6] или при фильтрации конфиденциальной информации в системах защиты от утечек данных [7].

Большинство пользователей глобальной сети Интернет ежедневно используют поисковые машины, но, как правило, текст, сравниваемый с массивом документов, имеет относительно ограниченную длину — максимум несколько десятков слов. Например, поисковый сервис Google ограничивает запрос 32 словами. Данная работа, относящаяся ко второму из указанных направлений, рассматривает задачи сравнения текстов произвольной длины.

Теоретически постановка задачи выглядит следующим образом: сравнить проверяемый текст t' произвольной длины с текстами t в массиве T , где размер массива T может быть достаточно большим, например, несколько миллионов документов. В определенных случаях такое сравнение может быть передано третьей стороне, обладающей высокопроизводительной компьютерной инфраструктурой для осуществления подобного поиска [8]. Однако следует заметить, что в ряде случаев задача поиска не может быть передана третьей стороне, так как сравниваемые тексты могут иметь конфиденциальный характер. В этом случае возникает проблема создания алгоритмов, позволяющих осуществлять поиск на горизонтально масштабируемой платформе, мощность которой может расти, следуя за потребностями организации. Подобная поисковая система должна обладать способностью к эластичному росту, не требующему существенных инвестиций в инфраструктуру, изменения алгоритмов и программного обеспечения.

Одним из возможных решений данной задачи является разработка структур данных и архитектуры поискового сервиса с использованием нереляционных баз данных, работающих на распределенной инфраструктуре [9].

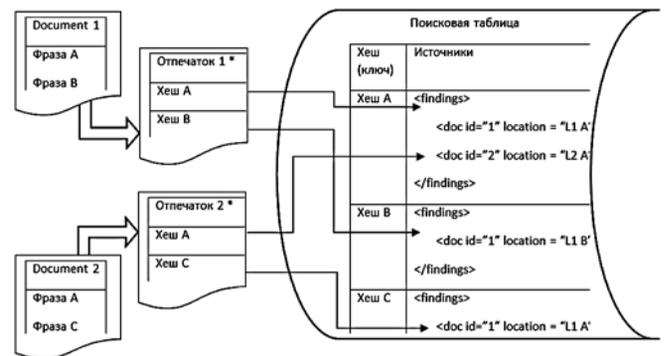
В данной статье рассмотрено подобное решение и показаны его преимущества и возможные ограничения на использование. Структурно работа построена следующим образом: в первом разделе дан краткий обзор существующих методов сравнения текстов с использованием традиционных баз данных [10, 11]. Во втором и третьем разделах рассмотрены представление данных для быстрого поиска и соответствующие эксперименты на больших текстовых выборках.

Модели хранения текстовых данных для быстрого сравнения

Большой класс методов быстрого текстового поиска по документам представлен методами, основанными на n -граммах (или шинглах). Данные методы предполагают создание отпечатка документа, который бы позволял быстро идентифицировать совпадающие части при попарном сравнении документов. За последние 30 лет разработаны десятки алгоритмов, основанных на шинглах. Подробный обзор данных алгоритмов с точки зрения производительности выполнен в работе [7]. Одним из известных алгоритмов, основанных на шинглах, является алгоритм *Winnowing* [12]. Пример применения *Winnowing*-подобного алгоритма для решения задачи сравнения текстов в системе подробнее рассмотрен в работе [13]. Показано, что выбранный шингл (n -грамма) помещается в таблицу базы данных в виде триады "`<хеш> — <документ> — <позиция хеша в документе>`" и представляется записями как минимум в двух реляционно-связанных таблицах. Примеры, приведенные в работах [13, 14], иллюстрируют высокую скорость роста базы данных шинглов при росте числа сравниваемых документов и, как следствие, скоростные ограничения, накладываемые реляционной моделью.

В качестве решения проблемы данная работа предлагает нереляционную модель типа "*ключ — значение*", где хеш подстроки (n -граммы) документа будет выступать в качестве ключа, а значение будет представлять собой двумерную таблицу, содержащую значения "`<документ> — <позиция хеша в документе>`" для всех документов, имеющих совпадающий текст. Таким образом, для поиска будет использоваться только одно ключевое значение, представленное значением хеш функции от n -граммы. Следствием этого является возможность распределения поиска по узлам кластера (шардам базы данных) для выполнения параллельных запросов к нескольким узлам одновременно [13].

Пример записи двух документов с совпадающим фрагментом представлен на рис. 1 [13]. В данном



* Для упрощения иллюстрации предположим, что в отпечаток попадают хеши, рассчитанные для каждой фразы

Рис. 1. Пример хранения данных в поисковой таблице [13]

примере в поисковую таблицу помещено два отпечатка документов с двумя совпадающими фрагментами (n -граммами).

Представленная теоретическая модель была реализована программно и протестирована в ряде экспериментов. Результаты тестирования рассмотрены ниже в следующем разделе.

Экспериментальные результаты

Эксперименты, выполненные в данной работе, направлены на подтверждение применимости использования распределенных баз данных для поиска схожих документов. Применимость оценивалась через два параметра — возможность масштабирования инфраструктуры и скорость работы системы. Скорость работы системы оценивалась во время загрузки базы сравнения в систему и во время сравнения одного документа с массивом текстов, загруженных в базу данных.

Возможность масштабирования инфраструктуры оценивалась через построение прототипа программной системы с использованием различных конфигураций горизонтально масштабированной базы данных. Были построены и протестированы инфраструктуры без разделения записей по индексу, а также с разделением записи по индексу между тремя и шестью шардами. При тестировании программы на всех трех инфраструктурах не требовалось изменять код приложения, что дает основание утверждать, что прототип системы может работать при разделении записей между фактически неограниченным числом узлов. Данное свойство обеспечивает возможность эластичного масштабирования системы при росте базы данных.

Схема кластера с тремя узлами хранения данных (шардами) показана на рис. 2 (см. третью сторону обложки). Кластер включает сервер приложения, работающий с сервером-маршрутизатором, который, в свою очередь, работает с конфигурационными серверами и с узлами хранения (шардами). В качестве физического оборудования для всех компонентов системы в экспериментах использовались неспециализированные персональные компьютеры офисного назначения.

В качестве базы данных сравнения была использована очищенная выборка статей из англоязычного сегмента Википедии (en.wikipedia.org). Документы в выборке были очищены от специфической разметки (тэгов) системы управления контентом MediaWiki (www.mediawiki.org). Исходя из предположения что документы, проверяемые на плагиат или на конфиденциальность, должны нести некую смысловую нагрузку, из полученного набора документов были удалены небольшие документы размером менее 100 слов. Полученная после подобной очистки выборка состояла из 3 917 453 документов, общий размер которых равен 9,84 Гбайт.

Оценка времени загрузки. Первый эксперимент имел целью оценить время загрузки исходного массива в базу данных. Данная операция не является критической для поисковой системы. В реальных условиях массовая загрузка большого числа документов выполняется при вводе системы в эксплуатацию, и в дальнейшем проводится рутинное обновление загруженного массива.

Среднее время загрузки (время/размер, мс/Кбайт) текста для различных конфигураций хранилища составляет: без шард — 87 мс/Кбайт, 3 шарда — 21 мс/Кбайт, 6 шардов — 11 мс/Кбайт. Зависимость времени загрузки от размера документа показана на графиках, приведенных на рис. 3—5.

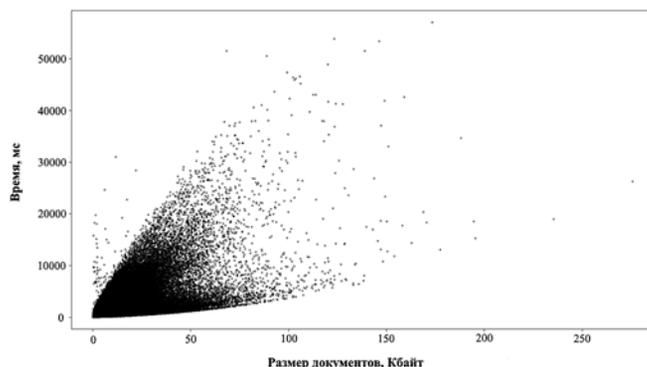


Рис. 3. Время записи в зависимости от размера документа (без шард)

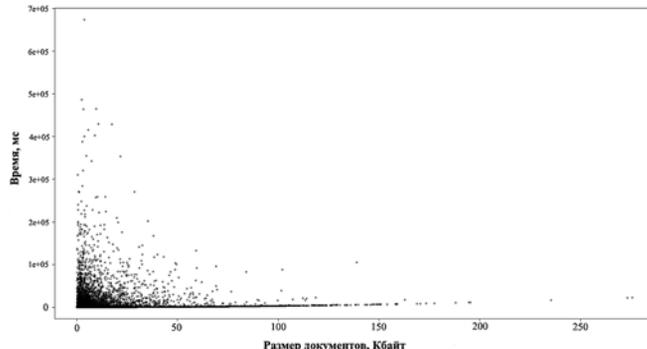


Рис. 4. Время записи в зависимости от размера документа (3 шарда)

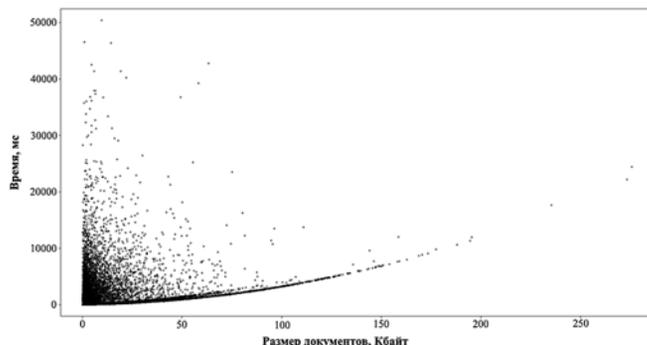


Рис. 5. Время записи в зависимости от размера документа (6 шард)

Однако, как показал анализ, данные графики не учитывают объем данных, уже загруженных в базу данных. На графиках явно видно, что время загрузки зависит не только от размера документа, так как значение времени значительно варьируется при загрузке документов приблизительно одного размера. Особенно хорошо это просматривается на графике, который находится на рис. 3.

С учетом того, что при загрузке проводится обновление индексов, размер базы данных является ключевым фактором, влияющим на производительность операции добавления. Графики, приведенные на рис. 6—8, показывают относительное время записи документов в зависимости от числа документов, уже загруженных в базу данных.

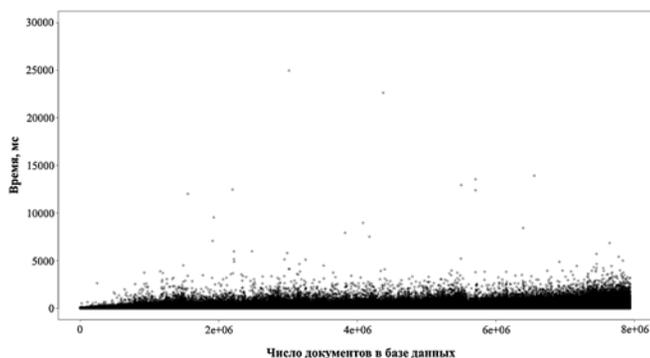


Рис. 6. Время поиска в зависимости от размера базы данных (без шард)

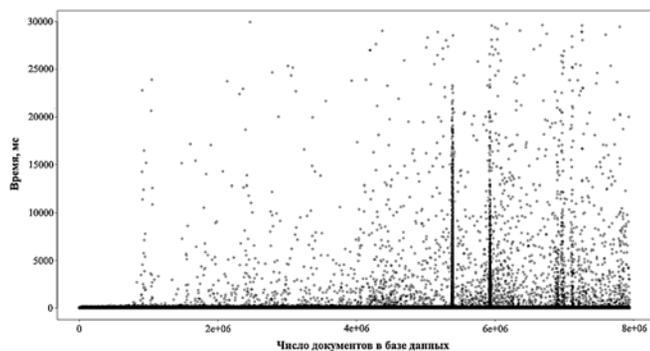


Рис. 7. Время поиска в зависимости от размера базы данных (3 шарда)

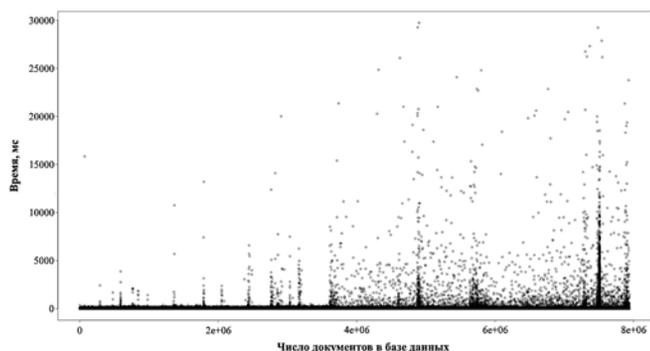


Рис. 8. Время поиска в зависимости от размера базы данных (6 шард)

Как видно из данных графиков, рост относительного времени, требуемого на загрузку в зависимости от объема уже загруженных данных, близок к линейному. При увеличении числа узлов хранения с одного до трех и с трех до шести заметно, что относительное время загрузки показывает меньший рост при увеличении числа узлов хранения. Это позволяет утверждать, что с точки зрения загрузки система является горизонтально масштабируемой.

На графиках, изображенных на рис. 7 и 8, отчетливо видно, что в момент перераспределения индекса по шардам время загрузки уменьшается. В случае с 3 шардами перераспределение происходит реже, но занимает больше времени по сравнению с работой кластера из 6 шард.

Оценка времени поиска. Второй эксперимент ставил целью оценку времени поиска по базе данных. Для проведения оценки массив из 3 917 453 документов был загружен в базу данных, установленную на загруженные ранее архитектуры. Для осуществления поиска из исходного массива были случайным образом отобраны 1000 документов и выполнен их поиск. Результатом каждого поиска было 100 % совпадения набора хешей документа, но эксперимент оценивал не качество, а время поиска. Графики времени поиска в зависимости от размера проверяемого документа приведены на рис. 9—11. Ниже представлено среднее время поиска на килобайт текста (время/размер, мс/Кбайт): без шард — 159,1 мс/Кбайт, 3 шарда — 77,7 мс/Кбайт, 6 шардов — 78,6 мс/Кбайт. Из этого можно сделать вывод, что среднее время без использования распределенной инфраструктуры как минимум в два раза выше, чем при хранении данных в распределенных узлах. Кроме того, график на рис. 9 показывает значительную нелинейность при росте объема проверяемого документа. В то же время графики на рис. 10 и 11 хорошо приближаются линейным представлением. Кроме того, предварительный анализ последних графиков показывает, что заметного роста времени поиска с увеличением числа шард не происходит, а даже наблюдается ее некоторое снижение при использовании 6 узлов вместо 3.

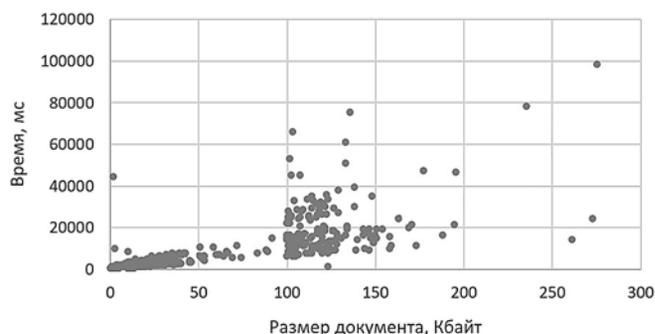


Рис. 9. Время поиска в зависимости от размера документа (без шард)

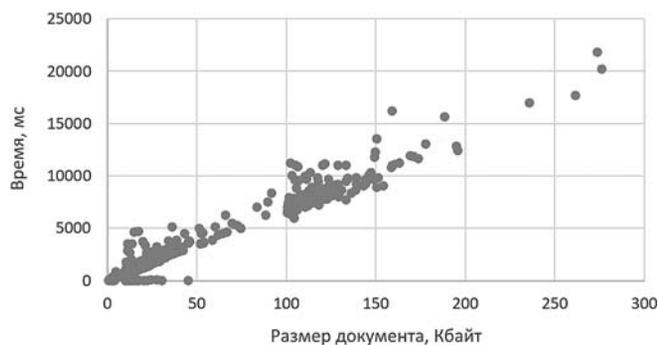


Рис. 10. Время поиска в зависимости от размера документа (3 шарда)

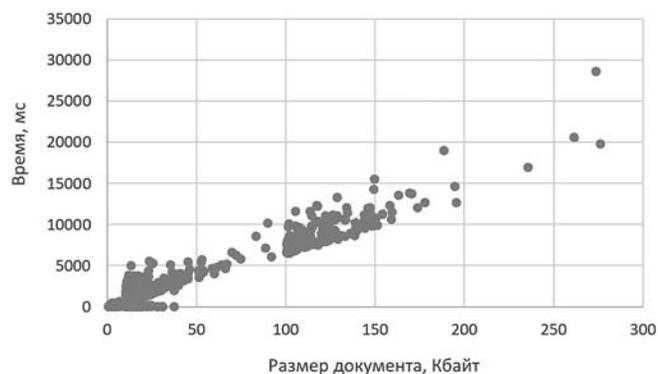


Рис. 11. Время поиска в зависимости от размера документа (6 шард)

Данный факт требует дополнительных исследований и может быть объяснен тем, что для данного объема проверяемых документов (~4 млн текстов) 6 узлов хранения значений отпечатка являются избыточными.

Практические выводы по результатам эксперимента и заключение по результатам работы представлены в следующем разделе.

Заключение

Количественная оценка времени поиска говорит о ее приемлемости для практических приложений, связанных с поиском плагиата или задачами защиты данных от утечек. Время 80 мс/Кбайт проверяемого текста может быть транслировано в односекундную задержку для текста длиной 6 страниц, что является приемлемым уровнем задержки для подобных систем. Если в качестве примера взять типичную задержку в 24 ч для систем проверки на плагиат, то можно утверждать, что пропускная способность подобной системы будет немного ниже 85 000 документов из 6 страниц в сутки, что является приемлемым для систем большого размера, например, обслуживающих один или несколько высших учебных заведений.

Следует отметить, что в данном случае система проводит лишь предварительный отбор документов

из большого массива. Алгоритмы, основанные на шинглах, не являются наилучшими с точки зрения детального сравнения документов. Следовательно, предварительный отбор должен быть дополнен детальным сравнением проверяемого документа и предварительно выбранных документов. С учетом того, что объем сравнения снизится со всей выборки в несколько миллионов документов до набора документов, которые гарантированно включают совпадающие фрагменты, детальное сравнение не должно вызвать значительных задержек во времени. Кроме того, процессы предварительного отбора и детального сравнения будут выполняться на различных подсистемах в параллельном режиме, что снизит их взаимозависимость.

Можно указать два направления дальнейших исследований: определение момента необходимости добавления узла при росте числа документов в базе данных и проведение эксперимента на большем числе узлов. Решение первой задачи возможно путем разработки модели оптимизации стоимости инфраструктуры при сохранении времени поиска. Эксперимент позволит верифицировать полученную модель на реальных данных.

Исследование выполнено при финансовой поддержке гранта совместной научной программы "РФФИ-Алтайский край", проект № 14-07-98000.

Список литературы

1. **Fomichov V. A.** K-calculus and K-languages as power formal means to design intelligent systems processing medical texts // *Cybernetica (Belgium)*. 1993. Vol. 36, N. 2. P. 161—182.
2. **Fomichov V. A.** Integral formal semantics and the design of legal full-text databases // *Cybernetica (Belgium)*. 1994. Vol. 37, N. 2. P. 145—177.
3. **Lei Y., Urea V., Motta E.** SemSearch: A Search Engine for the Semantic Web / Staab, Steffen, Svatek, Vojtech (Eds.) // *Proc. 15th International Conference on Knowledge Engineering and Knowledge Management "Managing Knowledge in a World of Networks"*, Podesbrady, Czech Republic, October 6—10, 2006, Proceedings. *Lect. Notes in Comp. Sci., Springer*, 2006. P. 238—245.
4. **Fomichov V. A., Kirillov A. V.** A Formal Model for Constructing Semantic Expansions of the Search Requests about the Achievements and Failures // *Artificial Intelligence: Methodology, Systems, and Applications, 15th International Conference, AIMSA 2012, Varna, Bulgaria, September 2012, Proceedings. Springer, Lecture Notes in Artificial Intelligence*. 2012. Vol. LNAI 7557. P. 296—304.
5. **HaCohen-Kerner Y., Tayeb A.** Experiments with Filtered Detection of Similar Academic Papers // *Artificial Intelligence: Methodology, Systems, and Applications, 15th International Conference, AIMSA 2012, Varna, Bulgaria, September 2012, Proceedings. Springer, Lecture Notes in Artificial Intelligence*. 2012. Vol. LNAI 7557. P. 1—13.
6. **Alzahrani S. M., Naomie S., Ajith A.** Understanding plagiarism linguistic patterns, textual features, and detection methods // *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions*. 2012. Vol. 42, N. 2. P. 133—149.
7. **Faro S., Lecroq Th.** The exact online string matching problem: A review of the most recent results // *ACM Computing Surveys*. 2013. Vol. 45, N. 13. P. 42—50. DOI=<http://dx.doi.org/10.1145/2431211.2431212>
8. **Демиденко Н. Д., Кулагин В. А., Шокин Ю. И.** Моделирование и вычислительные технологии распределенных систем. Новосибирск: Наука, 2012. С. 142—148.
9. **Торган Ю. Н., Зубрилина Т. В.** Использование нереляционного подхода в распределенной системе баз данных // *Науч-*

но-технические ведомости Санкт-Петербургского государственного политехнического университета. 2012. Т. 5. № 157. С. 15–20.

10. **Ganascia J. G., Lungo A. D.** Automatic detection of reuses and citations in literary texts // *Lit Linguist Computing*. 2014. Vol. 29, N. 3. P. 412–421.

11. **Pohuba D., Dulik T., Janku P.** Automatic evaluation of correctness and originality of source codes // In 10th European Workshop on Microelectronics Education (EWME). Tallinn, 2014. P. 49–52.

12. **Schleimer S., Wilkerson D., Aiken A.** Winnowing: Local Algorithms for Document Fingerprinting // *Proc. of the ACM SIG-*

MOD International Conference on Management of Data. San Diego, 2003. P. 76–85.

13. **Цхай А. А., Бутаков С. В., Мурзинцев С. В., Ким Л. С.** Обнаружение плагиата с использованием нереляционных баз данных // *Вестник алтайской науки*. 2015. № 1. С. 280–285.

14. **Дягилев В. В., Цхай А. А., Бутаков С. В.** Архитектура сервиса определения плагиата, исключающая возможность нарушения авторских прав // *Вестник Новосибирского государственного университета. Сер. "Информационные технологии"*. 2011. Т. 9. № 3. С. 23–29.

S. V. Butakov, Associate Professor, e-mail: sergey.butakov@computer.org,

S. V. Murzintsev, Postgraduate student, e-mail: o.l00@yandex.ru,

A. A. Tskhai, Full Professor, e-mail: taa1956@mail.ru,

Altai Academy of Economics and Law, Barnaul, Russia

Detecting Text Similarity on a Scalable Cluster

The paper addresses the problem of fast text comparison in massive datasets. Specific application areas in the project included plagiarism detection and text filtering in data loss protection systems. Survey of existing solutions based on relational databases in these areas outlined scalability limitations that may affect the performance. Solution proposed in this paper suggests using non-relational no-SQL databases with potential distribution of the search workload between nodes on the database cluster. To facilitate text search on distributed cluster, it was suggested to use "key-value"-like data structure for text representation. The work outlines details of the proposed data structure, describes developed software prototype and performed experiments. The latter confirmed applicability of the proposed solution on the distributed infrastructure with three and six nodes in terms of comparison quality and speed.

Keywords: text filtering, text comparison, big data, plagiarism detection, no-SQL, data loss prevention systems

References

1. **Fomichov V. A.** K-calculus and K-languages as power formal means to design intelligent systems processing medical texts <http://cat.inist.fr/?aModele=afficheN&cpsid=4211600>, *Cybernetica* (Belgium), 1993, vol. 36, no. 2, pp. 161–182.

2. **Fomichov V. A.** Integral formal semantics and the design of legal full-text databases, *Cybernetica* (Belgium), 1994, vol. 37, no. 2, pp. 145–177.

3. **Lei Y., Urea V., Motta E.** SemSearch: A Search Engine for the Semantic Web. Staab, Steffen, Svatek, Vojtech (Eds.), *Proc. 15th International Conference on Knowledge Engineering and Knowledge Management "Managing Knowledge in a World of Networks"*, Pödebrady, Czech Republic, October 6–10, 2006, Proceedings. Lect. Notes in Comp. Sci., Springer, 2006, pp. 238–245.

4. **Fomichov V. A., Kirillov A. V.** A Formal Model for Constructing Semantic Expansions of the Search Requests about the Achievements and Failures, *Artificial Intelligence: Methodology, Systems, and Applications, 15th International Conference, AIMS 2012*. Varna, Bulgaria, September 2012, Proceedings. Springer, Lecture Notes in Artificial Intelligence, 2012, vol. LNAI 7557, pp. 296–304.

5. **HaCohen-Kerner Y., Tayeb A.** Experiments with Filtered Detection of Similar Academic Papers, *Artificial Intelligence: Methodology, Systems, and Applications, 15th International Conference, AIMS 2012*. Varna, Bulgaria, September 2012, Proceedings. Springer, Lecture Notes in Artificial Intelligence, 2012, vol. LNAI 7557, pp. 1–13.

6. **Alzahrani S. M., Naomie S., Ajith A.** Understanding plagiarism linguistic patterns, textual features, and detection methods, *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions*, 2012, vol. 42, no. 2, pp. 133–149.

7. **Faro S., Lecroq Th.** The exact online string matching problem: A review of the most recent results, *ACM Computing Surveys*,

2013, vol. 45, no. 13, pp. 42–50, DOI=<http://dx.doi.org/10.1145/2431211.2431212>

8. **Demidenko N. D., Kulagin V. A., Shokin Ju. I.** *Modelirovanie i vychislitel'nye tehnologii raspredelennykh sistem: monografiya* [Modeling and computational technology distributed systems]. Novosibirsk: 2012, pp. 142–148.

9. **Torgan Ju. N., Zubrilina T. V.** Ispol'zovanie nerelacionnogo podhoda v raspredelennoy sisteme baz dannykh [Using the relational approach in a distributed database system], *Nauchno-tehnicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politehnicheskogo universiteta*, 2012, vol. 5, no. 157, pp. 15–20.

10. **Ganascia J. G., Lungo A. D.** Automatic detection of reuses and citations in literary texts, *Lit Linguist Computing*, 2014, vol. 29, no. 3, pp. 412–421.

11. **Pohuba D., Dulik T., Janku P.** Automatic evaluation of correctness and originality of source codes, *In 10th European Workshop on Microelectronics Education (EWME)*, Tallinn, 2014, pp. 49–52.

12. **Schleimer S., Wilkerson D., Aiken A.** Winnowing: Local Algorithms for Document Fingerprinting, *Proceedings of the ACM SIGMOD International Conference on Management of Data*, San Diego, 2003, pp. 76–85.

13. **Tskhai A. A., Butakov S. V., Murzincev S. V., Kim L. S.** Obnaruzhenie plagiata s ispol'zovaniem nerelacionnykh baz dannykh [Plagiarism detection using non-relational databases], *Vestnik altajskoj nauki*, 2015, no. 1, pp. 280–285.

14. **Djagilev V. V., Tskhai A. A., Butakov S. V.** Arhitektura servisa opredelenija plagiata, iskljuchajushhaja vozmozhnost' narusheniya avtorskih prav [The architecture of the service definition of plagiarism, which excludes the possibility of copyright infringement], *Vestnik Novosibirskogo gosudarstvennogo universiteta, serija "Informacionnye tehnologii"*, 2011, vol. 9, no. 3, pp. 23–29.

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ MODELING AND OPTIMIZATION

УДК 62-50:519.7/8

В. И. Левин, д-р техн. наук, проф., e-mail: vilevin@mail.ru,
Пензенский государственный технологический университет

Модели производной и их применение для моделирования динамических неполностью определенных систем

Рассмотрена проблема адекватности классического дифференциального исчисления. Указаны недостатки этого исчисления, связанные с неприменимостью классического понятия производной к неточно заданной функции. Введено понятие интервальной производной, применимое к функциям, задаваемым с точностью до интервала. Показано применение этого понятия к математическому моделированию неточно известных процессов.

Ключевые слова: интервал, интервальная функция, интервальная производная, интервальное дифференциальное исчисление, адекватность понятия производной

Введение

В XVII в. усилиями двух классиков естествознания — англичанина Исаака Ньютона и немца Готфрида Вильгельма Лейбница было открыто понятие производной функции и на его основе построено дифференциальное, а затем и интегральное исчисление [1]. Это открытие, позволившее формализованно изучать поведение самых разнообразных функций, имело колоссальное значение для последующего развития науки и расширения сферы ее приложений. С его помощью было получено огромное, не поддающееся точной оценке число новых научных результатов в самых разнообразных научных направлениях как в самой математике, так и во многих ее приложениях. Но время идет, подходы в науке меняются и, по нашему мнению, пришла пора разобраться, насколько модель производной Ньютона—Лейбница адекватна физической реальности и что надо сделать, чтобы эту адекватность улучшить. Решение проблемы позволит увеличить результативность исследований в различных областях науки и сделать их, с одной стороны, более глубокими и осмысленными, а с другой стороны, открывающими путь для более точного математического моделирования реальных систем и процессов.

1. Постановка проблемы

Как хорошо известно, понятие производной в классическом дифференциальном исчислении Ньютона—Лейбница [1] базируется на понятии предельного перехода, т. е. неограниченного приближения переменной величины — функции к некоторой постоянной величине — пределу, в процессе которого

разность между указанными величинами неограниченно приближается к нулю. Однако ясно, что такое предельное поведение возможно лишь в случае точно определенных функций и величин. На практике же при построении и исследовании математических моделей реальных систем и процессов различные их характеристики — функции и величины — всегда известны не полностью, а с некоторой ограниченной точностью (например, с точностью до интервалов возможных значений). В этих условиях никакого предельного перехода функции в смысле Ньютона—Лейбница не существует. Соответственно этому не существует ни классического понятия производной, ни классического дифференциального исчисления Ньютона—Лейбница. В связи со сказанным выше возникает, прежде всего, следующая базовая проблема: возможна ли переформулировка основных понятий классического дифференциального исчисления таким образом, чтобы они работали как в случаях, когда изучаемые функции и величины известны точно, так и в случаях, когда они известны с ограниченной точностью? Каким может быть это новое дифференциальное исчисление?

Вторая серьезная проблема в рассматриваемой нами области заключается в следующем. Как известно [1], производная в классическом дифференциальном исчислении имеет смысл скорости, с которой эта функция изменяется в различных точках относительно своего аргумента. В соответствии с этим производные высших порядков приобретают смысл скоростей высших порядков, с которыми данная функция изменяется относительно своего аргумента. Например, производная второго порядка означает ускорение — скорость изменения ско-

рости изменения нашей функции, т. е. скорость изменения функции второго порядка. Аналогично интерпретируются производные третьего и последующих порядков. Единообразная интерпретация производных различных порядков дает нам основание полагать, что производные всех порядков должны принадлежать одному и тому же классу функций. Однако на практике это условие в классическом дифференциальном исчислении не выполняется. Возникает естественный вопрос: означает ли эта разнородность производных различных порядков, что мы имеем дело с нормальным явлением, которое надо лишь правильно объяснить, или она свидетельствует о не вполне адекватной модели скорости изменения природных процессов, принятой в классическом дифференциальном исчислении, которую можно исправить, изменив модель и тем самым повысив точность моделирования различных систем и процессов? Какова эта новая модель?

Следующая проблема связана с тем, что скорость изменения природных процессов — это характеристика реально существующих процессов, которая поэтому должна обязательно существовать. Таким образом, любая функция, моделирующая реально существующий процесс, обязательно должна иметь всюду производную, если только само понятие производной введено адекватно. Но для обычной производной в классическом дифференциальном исчислении указанное требование, как известно, выполняется не всегда. В связи с этим уместен следующий вопрос: может ли модель скорости изменения реальных природных процессов, принятая в классическом дифференциальном исчислении, считаться вполне адекватной, учитывая, что в ней не всегда выполняется указанное естественное требование — существование скорости процесса, или же ее нужно признать не вполне адекватной и заменить другой, более адекватной моделью, в которой данное требование строго выполняется, что позволяет повысить точность моделирования разнообразных систем и процессов? Какова эта новая модель?

Далее, поскольку в классическом дифференциальном исчислении Ньютона—Лейбница производная любого порядка имеет один и тот же смысл скорости соответствующего порядка, с которой некоторая функция изменяется относительно своего аргумента, есть основание полагать, что производные всех порядков должны принадлежать одному и тому же классу функций (см. выше изложение второй проблемы). Из этого можно вывести, что должна существовать некоторая единая формула для производных всех порядков от любой заданной функции. Однако известно, что такой формулы в классическом дифференциальном исчислении нет. В связи с этим возникает следующий резонный вопрос: можно ли считать вполне адекватной модель скорости изменения природных процессов, принятую в классическом дифференциальном исчислении,

при том, что в ней не выполняется естественное требование — существование единообразной формулы для скорости (производной) любого порядка? Или же ее надо признать не вполне адекватной и заменить другой, более адекватной моделью, где выполняется указанное требование, что позволяет повысить точность моделирования различных систем и процессов? Какова тогда эта новая модель?

Наконец, последняя проблема в рассматриваемой области состоит в следующем. Понятие производной в любом из исчислений, в которых производная имеет смысл скорости изменения функции относительно своего аргумента, должно отвечать условию физичности в отношении размерности. Это требование связано с тем, что скорость — понятие физическое. Оно проявляется в том, что если, например, функция выражает зависимость пути в метрах от времени в секундах, то 1-я производная от этой функции по ее аргументу, выражающая скорость рассматриваемого движения, должна иметь размерность метр в секунду (м/с), 2-я производная, выражающая ускорение этого движения, т. е. скорость 2-го порядка, — размерность метр на секунду в квадрате (м/с^2) и т. д. Но в классическом дифференциальном исчислении это требование не всегда выполняется. Например, если зависимость пути в метрах s от времени в секундах t выражается функцией $s = e^t$, то скорость движения, выражаемая производной от s по t , имеет вид $s' = e^t$ и размерность, такую же, как s , т. е. метр, а не метр в секунду, как должно быть по условию физичности. В связи с этим возникает следующий вопрос: можно ли модель скорости изменения реальных процессов, принятую в классическом дифференциальном исчислении, считать вполне адекватной, хотя в ней не выполняется обязательное требование физичности в отношении размерности производной, или же надо признать не вполне адекватной и заменить другой, более адекватной моделью, в которой это требование выполняется, благодаря чему можно повысить точность моделирования различных систем и процессов? Какова может быть эта модель?

Новая модель производной, а также соответствующее ей дифференциальное исчисление, применимые для работы с функциями, заданными с ограниченной точностью, зависят от формы неопределенности, в которой задается функция. При стохастической неопределенности вводятся стохастическая производная и стохастическое дифференциальное исчисление (в частности, стохастические дифференциальные уравнения) [2]. Аналогично, в случае нечеткой неопределенности вводятся нечеткие производные, нечеткое дифференциальное исчисление, а также нечеткие дифференциальные уравнения [3—5] (обзор этой темы на русском языке см. в работах [6, 7]). Более простая и физичная версия всех этих понятий получается, если вместо нечетких функций использовать для введения неопреде-

ленности многозначные функции [8]. В настоящей работе проблема построения новой модели производной и соответствующего дифференциального исчисления рассматривается в предположении интервальной формы неопределенности функций. Это позволяет получать необходимые результаты в еще более простой и ясной форме, используя хорошо разработанный аппарат интервальной математики [9]. Кроме того, при таком подходе удастся конструктивно решить поставленные выше проблемы, вызванные неопределенностью изучаемых функций.

2. Решение базовых проблем

Рассмотрим сначала первую из проблем, поставленных в разд. 1, и укажем пути ее возможного решения.

Эта проблема, как уже указывалось, заключается в том, что при задании функций не точно, а с ограниченной точностью, понятие предела в смысле Ньютона—Лейбница, а вместе с ним и понятие производной перестают существовать. Возникает вопрос, можно ли распространить указанные понятия и на этот случай, чтобы получить возможность моделирования реальных процессов? Оказывается, что такое распространение возможно, если задавать функции с точностью до интервалов возможных значений, а в качестве математического аппарата использовать аппарат интервальной математики [9]. Следуя [10—12], используем понятие интервальной функции

$$\tilde{y} = \tilde{f}(\tilde{x}), \quad (1)$$

где $\tilde{x} = [x_1, x_2]$ — интервальная независимая переменная, выраженная в виде замкнутого вещественного интервала; $\tilde{y} = [y_1, y_2]$ — интервальная зависимая переменная, также выраженная в виде замкнутого вещественного интервала; \tilde{f} — однозначное отображение множества интервалов $\{\tilde{x}\}$ на множество интервалов $\{\tilde{y}\}$. Понятие предела интервальной функции (1) отличается от понятия предела обычной (точно заданной) функции. Пусть $\tilde{x} = [x_{01}, x_{02}]$ — интервальная постоянная. Будем говорить, что интервальная независимая переменная \tilde{x} в процессе своего изменения стремится к интервальному пределу \tilde{x}_0 , если x_1 стремится к пределу x_{01} , а x_2 — к пределу x_{02} . Символически это определение записывается в виде

$$\{\lim(\tilde{x} = [x_1, x_2]) = (\tilde{x}_0 = [x_{01}, x_{02}])\} \Leftrightarrow \{\lim x_1 = x_{01}, \lim x_2 = x_{02}\}. \quad (2)$$

Аналогично записывается определение предела интервальной зависимой переменной

$$\{\lim(\tilde{y} = [y_1, y_2]) = (\tilde{y}_0 = [y_{01}, y_{02}])\} \Leftrightarrow \{\lim y_1 = y_{01}, \lim y_2 = y_{02}\}. \quad (3)$$

Предел интервальной функции (1) теперь можем определить следующим образом. Пусть в про-

цессе своего изменения интервальная независимая переменная \tilde{x} стремится к интервальному пределу \tilde{x}_0 , при этом интервальная зависимая переменная \tilde{y} стремится к интервальному пределу \tilde{y}_0 . Будем в этом случае говорить, что предел интервальной функции (1) при \tilde{x} , стремящемся к \tilde{x}_0 , равен \tilde{y}_0 , или символически

$$\lim_{\tilde{x} \rightarrow \tilde{x}_0} \tilde{y} = \tilde{y}_0 \quad (4)$$

либо, по-другому, с использованием обозначения функции (1),

$$\lim_{\tilde{x} \rightarrow \tilde{x}_0} \tilde{f}(\tilde{x}) = \tilde{y}_0. \quad (5)$$

Теперь, используя понятие предела интервальной функции, мы можем ввести понятие производной интервальной функции (1). Будем полагать эту функцию непрерывной, т. е. ее нижнюю и верхнюю границы непрерывными функциями нижней и верхней границ интервала \tilde{x} — независимой переменной. Тогда фиксированному значению $\tilde{x}_0 = [x_{01}, x_{02}]$ независимой переменной будет соответствовать фиксированное значение функции $\tilde{y}_0 = \tilde{f}(\tilde{x}_0)$. Определим приращения независимой и зависимой переменных нашей функции относительно их указанных фиксированных значений:

$$\Delta\tilde{x} = \tilde{x} - \tilde{x}_0, \Delta\tilde{y} = \tilde{y} - \tilde{y}_0 = \tilde{f}(\tilde{x}) - \tilde{f}(\tilde{x}_0). \quad (6)$$

Составим отношение второго приращения к первому:

$$\begin{aligned} \Delta\tilde{y}/\Delta\tilde{x} &= (\tilde{y} - \tilde{y}_0)/(\tilde{x} - \tilde{x}_0) = \\ &= (\tilde{f}(\tilde{x}) - \tilde{f}(\tilde{x}_0))/(\tilde{x} - \tilde{x}_0). \end{aligned} \quad (7)$$

Возьмем предел отношения (7) при стремлении независимой переменной \tilde{x} к ее пределу \tilde{x}_0 :

$$\lim_{\tilde{x} \rightarrow \tilde{x}_0} \Delta\tilde{y}/\Delta\tilde{x}. \quad (8)$$

Предел (8), если он существует, назовем интервальной производной от интервальной функции (1) в точке \tilde{x}_0 и обозначим $\tilde{y}'_{\tilde{x}_0}$ или $\tilde{f}'_{\tilde{x}_0}(\tilde{x})$. Итак,

$$\tilde{y}'_{\tilde{x}_0} = \tilde{f}'_{\tilde{x}_0}(\tilde{x}) = \lim_{\tilde{x} \rightarrow \tilde{x}_0} \Delta\tilde{y}/\Delta\tilde{x}. \quad (9)$$

Доказано [10—12], что для существования в точке \tilde{x}_0 интервальной производной от интервальной непрерывной функции (1) необходимо и достаточно, чтобы в некоторой окрестности этой точки и в ней самой все значения независимой переменной \tilde{x} функции являлись невырожденными интервалами (интервалами с несовпадающими верхней и нижней границами). Также доказано [10—12], что интервальная производная от непрерывной интервальной функции в произвольной точке \tilde{x} может быть выражена в конечном виде через значения независимой переменной \tilde{x} и зависимой переменной

ной $\tilde{y} = \tilde{f}(\tilde{x})$ этой функции в указанной точке следующим образом:

$$\tilde{y}'_{\tilde{x}} \equiv \tilde{f}'_{\tilde{x}}(\tilde{x}) = (\tilde{f}(\tilde{x}) - \tilde{f}(\tilde{x})) / (\tilde{x} - \tilde{x}). \quad (10)$$

Отметим, что выражение (10) не следует воспринимать как неопределенность вида $0/0$, поскольку у любой существующей в точке \tilde{x} интервальной производной интервал \tilde{x} невырожден, и потому в соответствии с формулами интервальной математики знаменатель выражения (10) не равен нулю.

Формулы, приведенные в этом разделе, наглядно демонстрируют, что предпринятая переформулировка базовых понятий классического дифференциального исчисления — предела и производной — с использованием интервальной математики открывает возможность работы с этими понятиями в новых условиях, когда изучаемые величины и функции известны не точно, а с ограниченной точностью (в данном случае — с точностью до интервалов возможных значений). При этом предел сохраняет смысл постоянной величины, к которой неограниченно приближается переменная величина, а производная — смысл скорости, с которой указанная функция изменяется относительно своего аргумента. Здесь все постоянные и переменные величины, а также функции от них, являются интервальными. Эти новые математические понятия открывают возможность для адекватного моделирования реальных процессов, известных с ограниченной точностью.

Рассмотрим теперь вторую из проблем, поставленных в разд. 1. Эта проблема заключается в том, что, несмотря на единообразную интерпретацию в классическом дифференциальном исчислении Ньютона—Лейбница производных различных порядков как скоростей изменения функций различных порядков, в общем случае производные различных порядков не принадлежат одному и тому же классу функций. Является ли это положение нормальным, требующим лишь соответствующего объяснения, или же оно свидетельствует о некоторой неадекватности модели скорости изменения реальных процессов в классическом дифференциальном исчислении, которую, следовательно, необходимо исправить путем изменения модели? Обратимся к понятию интервальной производной, введенному выше. Она показывает скорость изменения интервальной функции $\tilde{y} = \tilde{f}(\tilde{x})$ относительно ее аргумента \tilde{x} . Запишем выражение интервальной производной (10) в виде

$$\tilde{y}' = \tilde{y}_{\Delta} / \tilde{x}_{\Delta}, \quad (11)$$

при следующих обозначениях:

$$\begin{aligned} \tilde{y}' &= \tilde{y}'_{\tilde{x}} = \tilde{f}'_{\tilde{x}}(\tilde{x}), \quad \tilde{y}_{\Delta} = \tilde{y} - \tilde{y} = \tilde{f}(\tilde{x}) - \tilde{f}(\tilde{x}), \\ \tilde{x}_{\Delta} &= \tilde{x} - \tilde{x}. \end{aligned} \quad (12)$$

Беря интервальную производную от интервальной производной функции 1-го порядка вида (11),

можно получить интервальную производную функцию 2-го порядка в виде

$$\tilde{y}'' = (\tilde{y}_{\Delta} / \tilde{x}_{\Delta})_{\Delta} / \tilde{x}_{\Delta}. \quad (13)$$

Аналогично, беря интервальную производную от интервальной производной 2-го порядка (13), получим интервальную производную 3-го порядка

$$\tilde{y}''' = ((\tilde{y}_{\Delta} / \tilde{x}_{\Delta})_{\Delta} / \tilde{x}_{\Delta})_{\Delta} / \tilde{x}_{\Delta} \quad (14)$$

и т. д. Выражение интервальной производной n -го порядка в общем виде имеет следующую форму:

$$\tilde{y}^{(n)} = \underbrace{(\dots}_{n-1 \text{ скобка}} \underbrace{(\tilde{y}_{\Delta} / \tilde{x}_{\Delta})_{\Delta} / \tilde{x}_{\Delta}}_{n-1 \text{ скобка}} \dots)_{\Delta} / \tilde{x}_{\Delta}. \quad (15)$$

Формулы интервальных производных (13)—(15) можно развернуть аналогично (12), используя выражение оператора Δ , указанное в (12). Как хорошо видно из приведенных формул (11)—(15), интервальные производные любого n -го порядка принадлежат одному и тому же классу интервальных функций, получаемых из исходной функции (1) по одному и тому же алгоритму с помощью лишь двух операций — деления и вычитания. Таким образом, можно считать, что понятие производной в интервально-дифференциальном исчислении изначально соответствует требуемой в постановке проблемы, рассмотренной в разд. 1, единообразной интерпретации производных различных порядков как скоростей различных порядков изменения функции. В то же время понятие производной в классическом дифференциальном исчислении не соответствует данной интерпретации, что свидетельствует о некоторой неадекватности модели скорости изменения реальных процессов, принятой в классическом дифференциальном исчислении. Так что переход от классического дифференциального исчисления к интервальному нужно рассматривать как некоторую корректировку модели производной в сторону ее большей адекватности процессам, происходящим в реальности, что способствует более точному моделированию указанных процессов с помощью интервально-дифференциального исчисления.

Третья из проблем, поставленных в разд. 1, состоит в том, что скорость изменения любых природных процессов, как характеристика реально существующих процессов, должна обязательно существовать. В связи с этим производная любой функции, описывающей природный процесс, должна обязательно существовать, если модель производной адекватна понятию скорости. Но на практике классическая производная существует не для всякой функции, описывающей природный процесс. Означает ли это некоторую неадекватность классической модели производной и можно ли эту неадекватность исправить путем перехода к новой модели? К какой?

Обратимся к понятию интервальной производной, введенному выше в разд. 2. Было указано, что

интервальная производная в любой точке непрерывной интервальной функции существует всегда, когда в данной точке и некоторой ее окрестности все значения независимой переменной этой функции являются невырожденными интервалами (т. е. интервалами с несовпадающими нижней и верхней границами). Таким образом, интервальная производная от интервальной функции существует практически всегда, в отличие от производной в классическом дифференциальном исчислении, которая во многих случаях не существует (например, в точках излома исходной функции). Снова видим, что переход от классической модели производной к интервальной можно считать корректировкой модели производной в сторону ее большей адекватности природным процессам, что дает возможность более точного моделирования этих процессов с помощью интервально-дифференциального исчисления.

Четвертая проблема, поставленная в разд. 1, состоит в том, что производная любого n -го порядка в классическом дифференциальном исчислении имеет смысл скорости изменения функции n -го порядка относительно своего аргумента (т. е. принципиально имеет тот же смысл — скорости изменения). Поэтому можно ожидать существования единой формулы для производных всех порядков от любой функции. Тем не менее в классическом дифференциальном исчислении такой формулы нет. Означает ли это некоторую неадекватность классической модели производной? Можно ли каким-нибудь образом скорректировать эту неадекватность, перейдя к новой модели производной? Снова обратимся к понятию интервальной производной. Рассмотрим формулу (15). Она дает единое выражение интервальных производных всех порядков от любой интервальной функции и тем самым полностью решает проблему. Таким образом, опять видим, что переход от классической модели производной к интервальной позволяет сделать ее более адекватной изучаемым реальным природным процессам и тем самым более точно моделировать эти процессы с помощью интервально-дифференциального исчисления.

Пятая проблема, поставленная в разд. 1, состоит, как уже было сказано, в том, что классическая модель производной не всегда удовлетворяет условию физичности в отношении размерности. Проявляется это в том, что скорости различных процессов, полученные путем вычисления соответствующих производных, не всегда имеют ту размерность, которую они должны иметь по законам физики. В связи с этим возникает некоторое сомнение в адекватности классической модели производной как скорости изменения реальных процессов и мыслить заменить ее другой, более адекватной моделью.

Обратимся к явным выражениям интервальных производных различных порядков — суперпозици-

ям интервалов $\tilde{y}_\Delta, \tilde{x}_\Delta$, (11), (13)—(15). Из выражений (12) интервалов \tilde{y}_Δ и \tilde{x}_Δ видно, что интервал \tilde{y}_Δ имеет размерность исходной функции $\tilde{y} = \tilde{f}(\tilde{x})$, от которой мы берем последовательные производные, а интервал \tilde{x}_Δ — размерность аргумента \tilde{x} этой функции. Обозначим эти размерности соответственно $\text{разм}\tilde{y}$ и $\text{разм}\tilde{x}$. Теперь, после применения формул (11), (13)—(15), легко получается, что размерность последовательных интервальных производных функции $\tilde{y} = \tilde{f}(\tilde{x})$ имеет вид

$$\begin{aligned} \text{разм}\tilde{y}' &= \text{разм}\tilde{y}/\text{разм}\tilde{x}, \\ \text{разм}\tilde{y}'' &= \text{разм}\tilde{y}/(\text{разм}\tilde{x})^2, \\ \text{разм}\tilde{y}''' &= \text{разм}\tilde{y}/(\text{разм}\tilde{x})^3, \dots, \text{разм}\tilde{y}^{(n)} = \\ &= \text{разм}\tilde{y}/(\text{разм}\tilde{x})^n. \end{aligned} \quad (16)$$

Но именно такой должна быть с точки зрения физики размерность этих производных, если исходить из их физического смысла как скоростей различных порядков изменения природных процессов. То есть получаемые в соответствии с формулами (16) размерности интервальных производных различных порядков полностью решают проблему. Например, если некоторая функция $\tilde{y} = \tilde{f}(\tilde{x})$ задает закон движения, в котором \tilde{x} означает время в секундах, \tilde{y} — пройденный путь в метрах, то 1-я производная \tilde{y}' , являясь скоростью движения, должна иметь размерность метр в секунду, вторая производная \tilde{y}'' , являясь ускорением, — размерность метр на секунду в квадрате ($\text{м}/\text{с}^2$) и т. д.

Снова видим, что переход от классической модели производной к интервальной делает модель более адекватной реальным процессам, тем самым позволяя более точно моделировать эти процессы с помощью математического аппарата интервально-дифференциального исчисления.

3. Алгоритм моделирования динамических неполностью определенных систем с использованием интервальных производных

Из представленных выше в разд. 2 результатов вытекает следующий алгоритм построения математической модели динамической неполностью определенной системы.

Шаг 1. Берем уже имеющуюся (или же строим заново) математическую модель динамической полностью определенной (детерминированной) системы, служащей прототипом динамической неполностью определенной (недетерминированной) системы, подлежащей моделированию (система-прототип получается из системы-оригинала путем детерминизации всех параметров системы, а система-оригинал получается из системы-прототипа путем их надлежущей раздетерминизации). Полагаем, что модель системы-прототипа имеет вид некоторого уравнения — алгебраического (если в нем есть только алгебраические операции), дифференциаль-

ного (если в дополнение к алгебраическим операциям в нем есть классические производные) или разностного (если в дополнение к алгебраическим операциям в нем фигурируют разностные операторы).

Шаг 2. Математическую модель системы-прототипа, полученную на шаге 1, подвергаем надлежущей раздетерминизации. Для этого заменяем все точно заданные параметры этой модели соответствующими интервальными параметрами, а все точно заданные операторы — соответствующими интервальными операторами (классическую производную — интервальной, обычный разностный оператор — интервальным). В результате получаем математическую модель системы-оригинала в символьной форме, где операторы (разностный и производный) представлены пока соответствующими символами, но интервальные параметры фигурируют уже в явном виде.

Шаг 3. Полученную на шаге 2 символьную математическую модель динамической неполностью определенной системы подвергаем необходимой десимволизации. Для этого заменяем все символы интервальных производных различных порядков соответствующими явными выражениями, применяя формулы (11), (13)—(15). С использованием аналогичных формул заменяем соответствующими явными выражениями все символы разностных операторов. В результате получаем математическую модель динамической неполностью определенной системы в явной интервальной форме, т.е. в форме суперпозиции интервальных параметров этой системы.

Конец процедуры

Полученная с помощью данного алгоритма математическая модель динамической неполностью определенной системы более адекватно отражает функционирование этой системы, чем детерминированная модель, полученная с помощью классического дифференциального исчисления.

Пример (демография). Известно, что скорость прироста населения прямо пропорциональна количеству населения. Найти модель, описывающую зависимость количества населения A от времени t , если известно, что в некоторый начальный момент $t = 0$ количество населения равнялось A_0 , а через год оно увеличилось на $a\%$. Построить математическую модель роста числа населения в условиях интервальной неопределенности системы.

Шаг 1. Скорость изменения количества населения есть 1-я производная от количества населения по времени, т.е. A' . Тогда по условию нашей задачи получаем дифференциальное уравнение — математическую модель роста количества населения $dA/dt = kA$ или, разделяя переменные,

$$dA/A = kdt.$$

Интегрируя почленно, находим

$$\ln A = \ln e^{kt + C_1},$$

что после потенцирования дает общее решение уравнения в виде

$$A = Ce^{kt}, C = e^{C_1}.$$

Сначала найдем константу C . Подставляя в общее решение соответствующие значения численности населения и времени, $A = A_0$, $t = 0$, получим

$$A_0 = Ce^{k \cdot 0} = C \cdot 1 = C,$$

так что общее решение принимает вид

$$A = A_0 e^{kt}.$$

Найдем константу k . По условиям задачи численность населения через год, т.е. при $t = 1$, составит $A = ((100 + a)A_0)/100$. Подставляя эти значения в общее решение, получим

$$((100 + a)A_0)/100 = A_0 e^k.$$

Откуда

$$e^k = (100 + a)/100,$$

$$k = \ln((100 + a)/100) = \ln\left(1 + \frac{a}{100}\right),$$

так что имеем следующее решение дифференциального уравнения — модели роста численности населения

$$A = A_0((100 + a)/100)^t.$$

Сама модель, представляющая основной интерес, приобретает окончательный вид следующего дифференциального уравнения:

$$dA/dt = \ln\left(1 + \frac{a}{100}\right)A, A|_{t=0} = A_0.$$

Шаг 2. Полученную на шаге 1 модель полностью определенной (детерминированной) демографической системы подвергаем раздетерминизации. Для этого точный параметр A мы заменяем интервальным параметром $\tilde{A} = [A_1, A_2]$, аналогично точные параметры A_0 , t — соответствующими интервальными параметрами $\tilde{A}_0 = [A_{01}, A_{02}]$, $\tilde{t} = [t_1, t_2]$. Классическую производную dA/dt заменяем интервальной производной $d\tilde{A}/d\tilde{t}$. В итоге получаем математическую модель демографической системы в виде интервально-дифференциального уравнения, записанного в символьной форме:

$$d\tilde{A}/d\tilde{t} = \ln\left(1 + \frac{a}{100}\right)\tilde{A}, \tilde{A}|_{\tilde{t}=[0, 0]} = \tilde{A}_0.$$

Шаг 3. Полученную на шаге 2 символьную математическую модель демографической системы десимволизируем, заменяя символьное выражение интервальной производной соответствующим явным интервальным выражением (11). В результате

получаем математическую модель демографической системы в явной интервальной форме

$$\frac{\tilde{A} - \tilde{A}}{\tilde{t} - \tilde{t}} = \ln\left(1 + \frac{a}{100}\right) \tilde{A}, \tilde{A}|_{\tilde{t} = [0, 0]} = \tilde{A}_0.$$

Заключение

Классическая производная и основанное на ней классическое дифференциальное исчисление были построены для исследования поведения точно определенных процессов. Однако в наше время большинство изучаемых процессов задаются неточно, так что их изучение с помощью классического дифференциального исчисления практически невозможно. Кроме того, оказывается, что классическое дифференциальное исчисление, даже при изучении точно определенных процессов, обнаруживает некоторые черты неадекватности по отношению к изучаемым процессам. Все это побуждает искать новые модели производных, пригодные для построения более адекватного дифференциального исчисления. Одной из них является интервальная производная, определяемая для функций, заданных с точностью до интервала возможных значений. Эта производная, предложенная ранее автором, детально изучена выше в целях установления ее адекватности реальным процессам. Наше изучение показало, что переход от классической производной к интервальной позволяет строить интервально-дифференциальное исчисление, которое более

адекватно процессам, заданным неточно, и таким образом позволяет точнее моделировать эти процессы. Кроме того, переход дает возможность избавиться от некоторых проявлений неадекватности, свойственных классической производной, и при изучении точно определенных процессов.

Список литературы

1. Фихтенгольд Г. М. Курс дифференциального и интегрального исчисления. Т. 1. М.: Физматлит, 2005. 650 с.
2. Леваков А. А. Стохастические дифференциальные уравнения. Минск: БГУ, 2009. 231 с.
3. Puri M. L., Ralescu D. A. Fuzzy Random Variables // Journal of Mathematical Application. 1986. N. 4. P. 409—422.
4. Friedman M., Ming M., Kandel A. Fuzzy Linear Systems // Fuzzy Sets and Systems. 1988. N. 96. P. 201—209.
5. Buckley J. J., Feuring J. Fuzzy Differential Equations // Fuzzy Sets and Systems. 2000. N. 110. P. 43—54.
6. Мочалов И. А., Хрисат М. С., Шихаб Еддин М. Я. Нечеткие дифференциальные уравнения в задачах управления. Ч. 1 // Информационные технологии. 2015. Т. 21, № 3. С. 171—178.
7. Мочалов И. А., Хрисат М. С., Шихаб Еддин М. Я. Нечеткие дифференциальные уравнения в задачах управления. Ч. 2 // Информационные технологии. 2015. Т. 21, № 4. С. 243—250.
8. Горбань И. И. Феномен статистической устойчивости. Киев: Наукова Думка, 2014. 370 с.
9. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. М.: Мир, 1987. 360 с.
10. Левин В. И. Интервальная производная и начала недетерминистского дифференциального исчисления // Онтология проектирования. 2013. № 4. С. 72—84.
11. Левин В. И. Интервально-дифференциальное исчисление и некоторые его применения // Информационные технологии. 2014. № 7. С. 3—10.
12. Левин В. И. Дифференциальное исчисление для интервально-определенных функций // Эвристические алгоритмы и распределенные вычисления. 2015. Т. 2, № 2. С. 8—25.

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

Models of Derivative and their Application to Modeling of Dynamical Incompletely Defined Systems

The problem of the adequacy of the classic differential calculus is considered in this article. Disadvantages of this calculation that prevent use it for simulation of real systems and processes in the uncertainty are discussed: inability to use classic concept of the derivative in the event of an incomplete initial functions; different functional accessory of classical derivatives of various orders, despite similar interpretation as the speed of various orders; nonexistence of the derivative in the individual points of functions that simulate the real movement despite the existence of the apparent rate at any point of each such process; the absence of a uniform formula for derivatives of various orders, despite a similar interpretation of all such derivatives as speeds of various orders; inconsistency of derivative to physicality in relation to dimension. In this regard, the problem of possibility of generalizing the classical differential calculus to circumvent all these shortcomings is discussed. It is shown how transition from classical derivative, which is introduced for exactly determined functions to interval derivative, which is incorporated to functions given up interval of possible values, allows to eliminate all the above drawbacks. Specifically, it is shown that the introduction of the concept of interval derivative allows you to build differential calculus for not fully certain original functions. It also allows you to standardize functional accessory derivatives of any order. Moreover, it provides for existence of the derivative of continuous interval functions. Also, a new calculus contains a single formula for derivatives of all orders and corresponds to the requirement of physicality in relation of the dimension. Algorithm of construction of interval models of dynamic incomplete defined systems using interval derivatives is given. An example of building of model of such system is presented.

Keywords: interval, interval function, function calculus, interval derivative, interval computing, nondeterministic differential calculus

References

1. **Fihtengol'c G. M.** *Kurs differencial'nogo i integral'nogo ischislenija*. Vol. 1. Moscow, Fizmatlit, 2005. 650 p. (in Russian).
2. **Levakov A. A.** *Stohasticheskie differencial'nye uravnenija*. Minsk: BGU, 2009. 231 p. (in Russian).
3. **Puri M. L., Ralescu D. A.** Fuzzy Random Variables, *Journal of Mathematical Application*, 1986, no. 4, pp. 409–422.
4. **Kandel A.** Fuzzy Linear Systems, *Fuzzy Sets and Systems*, 1988, no. 96, pp. 201–209.
5. **Buckley J. J., Feuring J.** Fuzzy Differential Equations, *Fuzzy Sets and Systems*. 2000, no. 110, pp. 43–54.
6. **Mochalov I. A., Hrisat M. S., Shihab Eddin M. Ja.** Nechetkie differencial'nye uravnenija v zadachah upravlenija. Part 1, *Informacionnye tehnologii*, 2015, vol. 21, no. 3. pp. 171–178 (in Russian).
7. **Mochalov I. A., Hrisat M. S., Shihab Eddin M. Ja.** Nechetkie differencial'nye uravnenija v zadachah upravlenija. Part 2, *Informacionnye tehnologii*, 2015, vol. 21, no. 4. pp. 243–250 (in Russian).
8. **Gorban' I. I.** *Fenomen statisticheskoj ustojchivosti*, Kiev, Naukova Dumka, 2014, 370 p. (in Russian).
9. **Alefel'd G., Hercberger Ju.** Vvedenie v interval'nye vychislenija, Moscow, Mir, 1987, 360 p.
10. **Levin V. I.** Interval'naja proizvodnaja i nachala nedeterministskogo differencial'nogo ischislenija, *Ontologija proektirovanija*, 2013, no. 4, pp. 72–84 (in Russian).
11. **Levin V. I.** Interval'no-differencial'noe ischislenie i nekotorye ego primenenija, *Informacionnye tehnologii*, 2014, no. 7, pp. 3–10 (in Russian).
12. **Levin V. I.** Differencial'noe ischislenie dlja interval'no-opredelennyh funkcij, *Jevristicheskie algoritmy i raspredelemnnye vychislenija*, 2015, vol. 2, no. 2, pp. 8–25 (in Russian).

УДК 004.94

Д. В. Исаев, канд. экон. наук, доц. кафедры бизнес-аналитики, disaev@hse.ru,
Национальный исследовательский университет "Высшая школа экономики", г. Москва

Имитационное моделирование сложных проектов

Рассмотрены вопросы имитационного моделирования сложных проектов, к числу которых относятся проекты с неоднозначным исходом, возможностью повторного выполнения и несколькими возможными вариантами реализации. Предложенные модели и рекомендации могут быть использованы для оценки сроков реализации и последствий программ развития, в состав которых входят сложные проекты. Практическое применение моделей проиллюстрировано на примере системы дискретно-событийного моделирования Arena.

Ключевые слова: программа развития, сложный проект, имитационное моделирование, дискретно-событийное моделирование, система Arena

Введение

Одной из важных особенностей программ развития сложных организационно-технических и социально-экономических систем является наличие в их составе проектов, характеристики которых имеют вероятностный характер. Это относится, прежде всего, к срокам выполнения отдельных проектов и их финансовым показателям. Также следует принимать во внимание наличие "сложных" проектов, к которым относятся проекты с неоднозначным исходом, проекты с возможностью повторного выполнения и проекты с несколькими возможными вариантами реализации. Наличие вероятностных характеристик во многом объясняет то, что в практике проектного менеджмента фактические показатели реализации проектов и программ крайне редко совпадают с запланированными.

Наличие вероятностных характеристик проектов свидетельствует о целесообразности имитационного моделирования реализации проектов и программ развития. При этом можно считать, что состояние объекта моделирования (начало и завершение проектов, наступление программных событий, сопутствующие платежи, получение результатов)

изменяется в дискретные моменты времени. Поэтому в данном случае уместно применение имитационного моделирования дискретно-событийного типа, которое дает возможность описывать изменение состояния исследуемой системы в виде последовательности событий, наступающих в дискретные моменты времени [1].

Таким образом, задачи данного исследования заключаются в выявлении основных типов и характеристик сложных проектов, имеющих место в программах развития организационно-технических и социально-экономических систем, а также в разработке практических рекомендаций в области их имитационного моделирования.

В качестве иллюстрации возможностей практического применения моделей рассматривается их реализация в среде информационной системы *Arena* (разработка компании *Rockwell Automation*), функциональность которой достаточно подробно описана как в зарубежной [2–4], так и в отечественной [5, 6] литературе. Одной из основных областей практического применения системы являются различные системы массового обслуживания, в том числе входящие в состав производственных систем и вычислительных комплексов. Систему *Arena* также

можно применять в управлении проектами. Например, в работе [7] описано ее использование для проектов со случайной продолжительностью и наличием случайных временных лагов между проектами. Данная модель получила развитие в авторской работе [8], где, помимо самих проектов, рассматривается их влияние на показатели зрелости объекта управления и соответствующие финансовые показатели. Дальнейшее развитие модели связано с рассмотрением проектов, имеющих сложную конфигурацию и в наибольшей степени подверженных воздействию фактора неопределенности.

Рассмотренные модели также могут быть реализованы в других информационных средах, поддерживающих методы дискретно-событийного моделирования. В частности, примерами российских разработок в данной области могут служить система *AnyLogic*, поддерживающая различные типы имитационных моделей, включая дискретно-событийный [9], и система *Actor Pilgrim*, предназначенная для моделирования временной, пространственной и финансовой динамики экономических процессов [10]. В качестве зарубежного аналога отметим систему *GPSS World* (разработка компании *Minuteman Software*), которая представляет собой мощную среду имитационного моделирования с использованием языка *GPSS*, хотя и не поддерживает графическую визуализацию моделей [11].

1. Постановка задачи имитационного моделирования проектов и программ развития

Общая постановка задачи имитационного моделирования программ развития и входящих в их состав проектов выглядит следующим образом.

Программа развития представляет собой совокупность проектов, выполняемых в определенной последовательности. Продолжительность каждого из проектов обычно является случайной. Дополнительное усложнение задачи возникает в следующих случаях:

- если результаты реализации проекта не могут быть заранее определены однозначно (проект с неоднозначным исходом);
- если имеется возможность повторного выполнения проекта в случае недостижения поставленных целей (проект с возможностью повторного выполнения);
- если проект может быть выполнен несколькими разными способами, в зависимости от конкретной ситуации, складывающейся на момент его начала (проект с несколькими вариантами реализации).

Все проекты, обладающие перечисленными свойствами, будем называть сложными проектами. При этом допускается наличие проектов с определенным сочетанием перечисленных выше факторов сложности.

Поскольку проекты выполняются в определенной последовательности, каждый из них может стартовать только после завершения всех предшествующих проектов (исключение составляют стартовые проекты программы развития, не имеющие предшествующих). Завершение нескольких проектов может трактоваться как наступление программного события. Могут иметь место задержки (как правило, случайные) между завершением предшествующих проектов и возможным началом последующих, а также между завершением проектов и наступлением последующих программных событий.

Следствием завершения проектов и наступления программных событий является изменение показателей зрелости объекта управления (системы, на совершенствование которой направлена программа развития), а также финансовых показателей, связанных с реализацией программы развития. Влияние проектов и событий на показатели может сопровождаться временными задержками (как правило, случайными).

Результатом имитационного моделирования являются такие характеристики программы развития, как сроки ее реализации, а также динамика показателей зрелости объекта управления и финансовых показателей программы.

2. Практическое моделирование проектов

Рассмотрим моделирование реализации проектов с применением программного обеспечения *Arena* как одного из наиболее распространенных инструментов, который, наряду с развитой функциональностью, обеспечивает наглядную визуализацию создаваемых моделей.

В системе *Arena* объект моделирования представляется в виде ориентированного графа, в качестве узлов которого выступают графические модули. Эти модули относятся к разным типам, при этом каждый тип имеет свое содержательное наполнение и графическое обозначение (именно эти обозначения использованы в последующих иллюстрациях). Также имеются модули данных, которые служат для описания общих параметров модели, но не участвуют в ее визуализации.

Изменение состояния объекта моделирования происходит в результате перемещений между узлами графа индивидуальных динамических объектов, называемых транзактами. В отличие от систем массового обслуживания при моделировании проектов транзакты являются абстрактными и не имеют физической интерпретации.

Рассмотрим на конкретном примере (рис. 1) использование графических модулей системы *Arena* для моделирования проектов, отношений предшествования между ними, а также их последствий. При этом будем рассматривать только ту часть функциональности системы, которая является задействованной для решения сформулированной задачи.

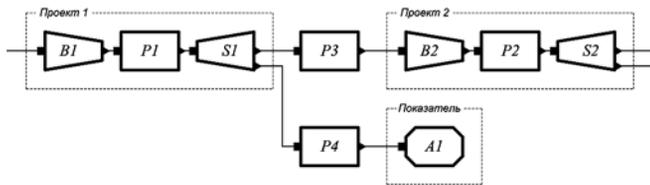


Рис. 1. Проекты, отношения предшествования и влияние на показатели

Основным при описании проекта является графический модуль типа **Process** (на рис. 1 модули этого типа имеют обозначения *P1* и *P2*). Поступление в этот модуль транзакта означает начало проекта, а выход транзакта из модуля — завершение проекта. В каждом из модулей данного типа указывается задержка — продолжительность реализации проекта. Период задержки может быть как детерминированным, так и случайным, в последнем случае задаются распределение случайной величины и его параметры.

При моделировании предполагается, что программа развития является стратегически важной для компании, и любой из проектов обеспечен достаточным объемом ресурсов — либо внутренних, либо привлекаемых со стороны. Поэтому конкретные ресурсы, как и уровень загрузки мощностей (актуальные для производственных систем), в данном случае не рассматриваем, вместо этого принимаем во внимание агрегированные финансовые показатели программы развития.

Поскольку проекты связаны между собой отношениями предшествования, каждый из них может иметь несколько предшествующих проектов и несколько последующих. Для описания таких отношений используют графические модули типов **Batch** и **Separate**.

Модуль **Batch** позволяет описать правило, согласно которому проект может стартовать только после завершения всех предшествующих проектов (на рис. 1 такие модули имеют обозначения *B1* и *B2*). В модуле осуществляется группировка входящих транзактов, при этом параметр объема группировки задается равным числу предшествующих проектов. В результате на входе модуля **Batch** возникает очередь из транзактов, ожидающих группировки. Сама группировка происходит в момент поступления на вход модуля последнего из ожидаемых транзактов, что означает завершение последнего из предшествующих проектов. Результатом группировки является единственный транзакт, который передается из модуля **Batch** в модуль **Process**. Если имеется лишь один предшествующий проект (или если проект является стартовым в программе развития), то модуль **Batch** можно не использовать.

Влияние рассматриваемого проекта на последующие описывается с помощью модуля типа **Separate** (на рис. 1 модули этого типа обозначены *S1* и *S2*).

На вход модуля поступает единственный транзакт (из модуля **Process**). В модуле **Separate** этот транзакт тиражируется, в результате чего возникают два новых одинаковых транзакта, один из которых условно называется оригиналом, а другой — копией. Таким образом, модуль **Separate** позволяет описать связь проекта с двумя последующими объектами. Если число последующих объектов больше двух, то используется "каскад" из нескольких модулей этого типа. Если же проект имеет лишь один последующий элемент (проект, событие или показатель), то модуль **Separate** не используется.

Для описания временных лагов, отделяющих завершение одних проектов от возможного начала других, используют такие же модули, как и те, которые применяют для описания продолжительности проектов, а именно — модули типа **Process**, с указанием детерминированных или случайных задержек. В рассматриваемом примере (см. рис. 1) проект 1 предшествует проекту 2, а разделяющий их временной лаг характеризует модуль *P3*.

По аналогии с проектами осуществляют моделирование программных событий, с той разницей, что период задержки, указываемый в модуле **Process**, задают равным нулю.

Для описания показателей зрелости объекта управления и финансовых показателей, связанных с реализацией программы развития, используют глобальные переменные (глобальными их называют потому, что они относятся не к какому-либо отдельному модулю, а к модели в целом). Глобальные переменные могут представлять собой не только отдельные переменные, но и массивы — одномерные или двумерные. В частности, для значений показателей зрелости используют одномерный массив с числом элементов, равным числу показателей. Аналогичный одномерный массив используют для финансовых показателей.

Начальные значения элементов массивов показателей задают в модуле данных **Variable** (этот модуль не является графическим и не используется при визуализации модели). В процессе моделирования изменения элементов массивов описываются с помощью графических модулей типа **Assign**. В рассматриваемом примере (см. рис. 1) проект 1 влияет на изменение показателя, а соответствующий модуль **Assign** имеет обозначение *A1*.

Степень влияния проектов и событий на показатели также описывается с помощью глобальных переменных, представляющих собой одномерные массивы. Значение каждого из элементов массива, описывающего воздействия на показатели зрелости, представляет собой гарантированное значение показателя, которое он получает в результате реализации соответствующего проекта или наступления соответствующего события. Аналогично, значение элемента массива, описывающего финансовые последствия, представляет собой исходящий денежный

поток, связанный с соответствующим проектом или событием. Значения элементов массивов воздействий также заданы в модуле данных **Variable**.

Как и в случае с предшествованием проектов, могут иметь место временные лаги между завершением проектов (или наступлением программных событий) и изменением показателей. Такие задержки также описывают с помощью модулей типа **Process** (на рис. 1 такой модуль имеет обозначение *P4*).

Наконец, рассмотрим элементы, задействованные при начале и завершении сеанса моделирования.

Начало сеанса моделирования описывается с помощью единственного графического модуля типа **Create**. В момент начала процесса моделирования этот модуль генерирует единственный транзакт, который подлежит тиражированию с использованием модулей **Separate** (одного или нескольких, в зависимости от объема тиражирования), с последующей доставкой возникших в результате тиражирования транзактов на вход стартовых проектов программы развития (т. е. проектов, не имеющих предшествующих). Возможные задержки между началом моделирования и возможными началами стартовых проектов описывают с помощью модулей **Process**.

Завершение сеанса моделирования описывается графическим модулем **Dispose**, который терминирует поступающие в него транзакты. Сеанс моделирования считается завершенным после терминирования последнего из транзактов, имеющихся в системе.

Приведенная модель позволяет описывать программы развития, включающие "простые" проекты, т. е. проекты, не допускающие вариативности получаемых результатов, возможности повторов, а также неопределенности в части способов их выполнения. Однако проекты с перечисленными свойствами ("сложные" проекты) также важны, поскольку довольно часто встречаются при развитии социально-экономических и организационно-технических систем. Поэтому рассмотрим основные типы сложных проектов, их роль в программах развития, а также возможности их моделирования с помощью системы *Arena*.

3. Проекты с неоднозначным исходом

Одной из разновидностей сложных проектов является проект с неоднозначным исходом. Необходимость рассмотрения таких проектов объясняется тем, что довольно часто при планировании проекта бывает трудно оценить не только его продолжительность, но и результаты, получаемые в ходе его выполнения. Результаты (исходы) проекта оказывают существенное влияние на его последствия — выполнение последующих проектов, наступление последующих проектных событий, изменение значений показателей зрелости объекта управления и финансовых показателей. Поэтому можно

считать, что каждому из возможных исходов проекта соответствует свой набор последствий.

Примерами проектов с неоднозначным исходом могут служить научно-исследовательские работы, в которых от результатов некоторых экспериментов зависит дальнейший ход исследования. Другим примером могут служить геолого-разведочные работы: здесь стратегия разведки также будет меняться в зависимости от результатов, полученных на предшествующих этапах.

При моделировании проектов с неоднозначным исходом может быть использован вероятностный подход. При этом считается, что каждый из возможных результатов реализации проекта достигается с некоторой вероятностью, сумма которых по всем возможным исходам равна единице.

Рассмотрим моделирование таких проектов в системе *Arena* на примере проекта с двумя возможными исходами (рис. 2).

Как и в случае простого проекта, модель сложного проекта с неоднозначным исходом имеет модуль **Process** с заданной продолжительностью выполнения проекта и модуль **Batch** с объемом группировки, равным числу предшествующих проектов (на рис. 2 эти модули имеют обозначения *P1* и *B1*). Однако в отличие от простого проекта, в данном случае предусматривается возможность получения одного из двух альтернативных результатов. Эта неопределенность описывается с помощью графического модуля **Decide**, который имеет обозначение *D1*.

По сути, модуль **Decide** характеризует ветвление процесса, при этом могут быть предусмотрены как две исходящие ветви, так и несколько исходящих ветвей (множественный выбор). В модели проекта число исходящих ветвей равно числу возможных исходов (в рассматриваемом на рис. 2 случае — двум). Выбор той или иной ветви для маршрутизации транзакта может осуществляться либо случайным образом, либо в соответствии с заданным условием. В случае вероятностного выбора (что предусмотрено постановкой рассматриваемой задачи) в качестве параметров модуля **Decide** задаются вероятности, в соответствии с которыми транзакт, поступивший в модуль, будет направлен по той или иной исходящей ветви (при этом сумма вероятностей по всем исходящим ветвям должна быть равна единице).

Отличие от простого проекта состоит также и в том, что в модели проекта с неоднозначным исходом предусматривается не один, а несколько модулей

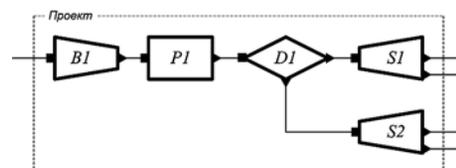


Рис. 2. Проект с неоднозначным исходом

Separate. Число таких модулей равно числу возможных исходов и, соответственно, числу выходов модуля **Decide** (в рассматриваемом примере с двумя возможными исходами имеются два модуля **Separate** с обозначениями *S1* и *S2*).

В модели модуль **Decide** расположен между модулем **Process** и модулями **Separate**. Транзакт, покидающий модуль **Process** при завершении проекта, поступает на вход модуля **Decide**. Здесь происходит его маршрутизация: в соответствии с заданными вероятностями транзакт покидает модуль **Decide** по той или иной исходящей ветви и, соответственно, поступает на вход того или иного модуля **Separate**. Дальнейшее поведение системы определяется дугами, связывающими модуль **Separate** (или каскад модулей данного типа, если число последующих элементов больше двух) с последующими модельными объектами — проектами, событиями и показателями.

4. Проекты с возможностью повторного выполнения

Еще одним типом сложных проектов является проект с возможностью повторного выполнения. В этом случае проект считается полностью завершенным только тогда, когда его цели полностью достигнуты. Однако, если цели проекта не достигнуты, то есть возможность выполнить проект заново. Необходимость повторного выполнения проекта носит вероятностный характер: с одной вероятностью цели проекта достигаются и проект считается полностью завершенным, с другой вероятностью цели не достигаются, и поэтому требуется его повторное выполнение (сумма двух вероятностей равна единице).

Также могут возникать ситуации, когда повторное выполнение проекта возможно только по истечении некоторого промежутка времени (детерминированного или случайного) после завершения предыдущей итерации.

Проекты с возможностью повторного выполнения могут возникать при создании инновационных продуктов или внедрении принципиально новых (еще не апробированных) технологических решений. В этом случае успех проекта не гарантирован, и в случае неудачи проект инициируется заново.

Пример проекта с возможностью повторного выполнения приведен на рис. 3.

Как и для простого проекта, модель проекта с возможностью повторного выполнения предусматривает модуль **Process** с заданной продолжительностью

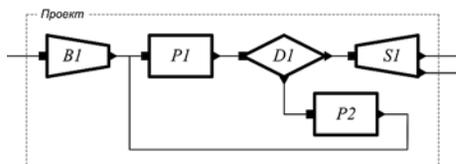


Рис. 3. Проект с возможностью повторного выполнения

стью выполнения проекта и модуль **Batch**, у которого параметр объема группировки равен числу предшествующих проектов (на рис. 3 эти модули обозначены *P1* и *B1*). Как и в случае проекта с неоднозначным исходом, присутствует графический модуль **Decide**, имеющий обозначение *D1*. Однако отличие состоит в том, что модуль **Decide** всегда имеет лишь две ветви, одна из которых характеризует достижение цели проекта и, соответственно, его окончательное завершение, а другая — необходимость повторного выполнения проекта. Транзакт, поступивший на вход модуля **Decide**, направляется по одной из двух исходящих ветвей. Выбор ветви для маршрутизации осуществляется в соответствии с заданными в модуле **Decide** вероятностями, сумма которых равна единице.

В отличие от проектов с неоднозначным исходом, для проекта с возможностью повторного выполнения предусмотрен лишь один модуль **Separate**. Транзакт поступает на вход этого модуля только в случае, когда цели проекта считаются достигнутыми и его повторное выполнение не требуется.

Как и в случае проекта с неоднозначным исходом, модуль **Decide** расположен между модулями **Process** и **Separate**. Один из выходов модуля **Decide** (соответствующий успешному выполнению проекта) соединяется со входом модуля **Separate**. Другой выход (характеризующий необходимость повторного выполнения) соединяется со входом модуля **Process**. В случае если имеет место временной лаг между завершением проекта и началом его повторного выполнения, используется дополнительный модуль **Process**, в котором описывается соответствующая задержка (на рис. 3 этот модуль имеет обозначение *P2*).

Транзакт, покидающий модуль **Process** при завершении очередной итерации проекта, поступает на вход модуля **Decide**, где происходит его маршрутизация. Если проект признается успешным, то транзакт перенаправляется на вход модуля **Separate**, а в случае необходимости повторного выполнения проекта — сначала в модуль **Process**, характеризующий задержку начала повторного выполнения, а затем — в модуль **Process**, описывающий непосредственную реализацию проекта.

5. Проекты с несколькими вариантами реализации

К числу сложных также относятся проекты, допускающие несколько вариантов их реализации. Рассмотрение таких проектов можно считать целесообразным (а в некоторых случаях — даже необходимым), когда цели проекта могут быть достигнуты разными способами (например, с применением разных технологий или организационных схем), но выбор конкретного способа зависит от ситуации, имеющей место на момент начала проекта. На стадии формирования потенциальной программы развития можно определить совокуп-

ность возможных ситуаций и разработать для каждой из них свой вариант реализации проекта. Также могут быть сделаны предположения относительно вероятности появления каждой из ситуаций. Тогда конкретный способ реализации проекта будет определен в момент его начала с той или иной вероятностью в зависимости от складывающихся обстоятельств (сумма таких вероятностей по всем возможным ситуациям, связанным с рассматриваемым проектом, равна единице).

Примером проекта с несколькими вариантами реализации является внедрение информационной системы управления в случае, если применение того или иного программного обеспечения соответствующего класса будет определяться некоторыми объективными условиями (например, перспективами выхода компании-разработчика на российский рынок). В этом случае решение о выборе программного обеспечения (и соответственно, параметры проекта, характерные для этой альтернативы) носит условный характер, а окончательный выбор осуществляется при начале проекта, в зависимости от сложившейся на тот момент ситуации.

Пример проекта с несколькими возможными вариантами реализации приведен на рис. 4.

Как и в предыдущих случаях, в состав модели проекта с несколькими вариантами реализации входит единственный модуль **Batch**, с объемом группировки, равным числу предшествующих проектов (на рис. 4 этот модуль имеет обозначение *B1*). Также присутствует графический модуль **Decide**, у которого могут быть как две, так и несколько исходящих ветвей (в рассматриваемом примере этот модуль обозначен *D1*). Число исходящих ветвей модуля **Decide** равно числу вариантов реализации проекта. Выбор ветви для маршрутизации, соответствующей тому или иному варианту реализации проекта, проводится в соответствии с заданными вероятностями, сумма которых равна единице.

Транзакт, поступивший на вход модуля **Decide**, впоследствии направляется по одной из нескольких исходящих ветвей. При этом каждой из исходящих ветвей соответствует по одному модулю **Process** и одному модулю **Separate**, которые последовательно соединены друг с другом (в рассматриваемом примере для первой ветви эти модули имеют обозначения *P1* и *S1*, а для второй — *P2* и *S2*).

Для каждого из модулей **Process** указываются свои значения параметров, в частности продолжительность проекта, которая может быть как детер-

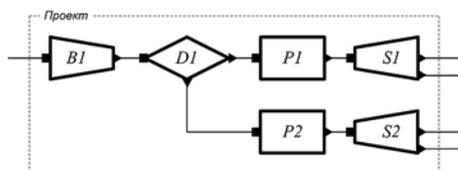


Рис. 4. Проект с несколькими вариантами реализации

минированной, так и случайной. Этим проект с несколькими вариантами реализации отличается от проекта с неоднозначным исходом и проекта с возможностью повторного выполнения, где реализация проекта осуществляется в соответствии с единственным набором параметров, заданных в модуле **Process**.

В модели проекта с несколькими вариантами реализации транзакт, покидающий модуль **Batch** в момент начала проекта, поступает на вход модуля **Decide**. В модуле **Decide** осуществляется маршрутизация, определяющая выбор конкретного варианта реализации проекта. В зависимости от сделанного выбора, транзакт передается в тот или иной модуль **Process**, характеризующий соответствующий вариант реализации проекта и имеющий свой набор параметров. Наконец, после модуля **Process** транзакт поступает на вход соответствующего модуля **Separate**, после чего дальнейшее поведение системы определяется исходящими дугами модуля **Separate** (или каскада модулей этого типа, если число последующих модельных объектов больше двух), а также самими последующими модельными объектами.

Заключение

Многие из проектов, входящих в состав программ развития сложных организационно-технических и социально-экономических систем, находятся под значительным влиянием случайных факторов. Помимо вероятностного характера продолжительности выполнения проектов (что характерно для большинства из них), можно выделить такие факторы, как неоднозначность ожидаемых результатов, возможность повторного выполнения проекта, а также возможность реализации проекта несколькими разными способами. Перечисленные факторы могут сочетаться друг с другом, например, может рассматриваться проект с несколькими вариантами реализации, каждый из которых допускает неоднозначность получаемых результатов, в том числе приводящих к необходимости повторного выполнения проекта.

Перечисленные факторы существенно усложняют планирование проектов и программ развития, а также оценку последствий их реализации. Поэтому, принимая во внимание дискретный характер происходящих изменений, представляется уместным применение дискретно-событийного имитационного моделирования.

Одним из апробированных и хорошо зарекомендовавших себя инструментов дискретно-событийного моделирования является система *Arena*. Ее важными достоинствами являются возможность гибкого описания различных задержек (в том числе случайных), а также численных значений показателей, характеризующих ресурсоемкость и результативность проектов. Что касается сложных проектов (с неоднозначным исходом, возможностью повтор-

ного выполнения и наличием нескольких вариантов реализации), то значительную роль в их моделировании играют возможности вероятностной маршрутизации, которые в системе *Arena* обеспечивает специальный модуль выбора (модуль типа **Decide**). Таким образом, функциональность системы *Arena* является вполне достаточной для моделирования даже самых сложных проектов.

Практическая значимость предложенного подхода состоит в том, что на основе многократного прогона модели (возможно с разными наборами параметров) можно оценивать ключевые характеристики программ развития, включая сроки реализации проектов и наступления ключевых событий, а также показатели эффективности программ, определяемые на основе соотношения ожидаемых затрат и результатов. Это свидетельствует о применимости предложенной модели на стадии разработки и планирования программ развития, а при наличии нескольких альтернативных программ ее использования также и для принятия решения о выборе одной из них для реализации.

Список литературы

1. **Wainer G. A.** *Discrete-event modeling and simulation: A practitioner's approach*. Boca Raton, FL: CRC Press, 2009. 520 p.
2. **Altiock T., Melamed B.** *Simulation modeling and analysis with Arena*. San Diego: Academic Press, 2007. 456 p.
3. **Kelton W. D., Sadowski R. P., Swets N. B.** *Simulation with Arena*. London: McGraw-Hill, 2014. 656 p.
4. **Rossetti M. D.** *Simulation modeling and Arena*. New York: John Wiley & Sons, 2015. 744 p.
5. **Замятина О. М.** *Моделирование систем*. Томск: ТПУ, 2009. 204 с.
6. **Щербakov С. М.** *Имитационное моделирование экономических процессов в системе Arena*. Ростов н/Д: РИНХ, 2012. 128 с.
7. **Cosgrove W. J.** *Simplifying PERT network simulation with Arena* // *California Journal of Operations Management*. 2008. Vol. 6, N. 1. P. 61—68.
8. **Исаев Д. В.** *Анализ программ развития систем управления с применением программного обеспечения Arena* // *Аудит и финансовый анализ*. 2015. № 6. С. 442—447.
9. **Бунцев И. А.** *Создание и реализация имитационных моделей в программной среде AnyLogic*. М.: Горячая линия — Телеком, 2015. 153 с.
10. **Емельянов А. А., Емельянова Н. З.** *Имитационное моделирование и компьютерный анализ экономических процессов*. Смоленск: Универсум, 2013. 266 с.
11. **Бражник А. Н.** *Имитационное моделирование: возможности GPSS WORLD*. СПб.: Реноме, 2006. 439 с.

D. V. Isaev, Associate Professor, e-mail: disaev@hse.ru,

Department of Business Analytics National Research University Higher School of Economics

Simulation of Complicated Projects

The paper focuses on simulation of complicated projects, associated with developing of socio-economic, organizational and technical systems. Few types of complicated projects are determined: projects with few possible outcomes, repeatable projects, and projects with few possible ways of implementation. Combined cases with certain combination of the factors of complexity may also be considered.

The purpose of the research is to determine key characteristics of complicated projects, and to develop recommendations in the field of their discrete-event simulation.

For illustration of practical application of the models, Arena software, as a popular and approved tool of discrete-event simulation, is applied. A set of models for different types of complicated projects are presented. Each of the models include detailed description of appropriate Arena modules, their functions in the model and their relevant parameters. The models and related recommendations advanced in the paper are applicable for assessment of implementation terms and consequences of development programs with complicated projects.

Keywords: *development program, complicated project, simulation, discrete-event modeling, Arena software*

References

1. **Wainer G. A.** *Discrete-event modeling and simulation: A practitioner's approach*, Boca Raton, FL, CRC Press, 2009, 520 p.
2. **Altiock T., Melamed B.** *Simulation modeling and analysis with Arena*, San Diego, Academic Press, 2007, 456 p.
3. **Kelton W. D., Sadowski R. P., Swets N. B.** *Simulation with Arena*, London, McGraw-Hill, 2014, 656 p.
4. **Rossetti M. D.** *Simulation modeling and Arena*, New York: John Wiley & Sons, 2015, 744 p.
5. **Zamyatina O. M.** *Modelirovanie sistem* [Modeling of systems], Tomsk, Tomsk Polytechnic University, 2009, 204 p.
6. **Shcherbakov S. M.** *Imitatsionnoe modelirovanie ekonomicheskikh protsessov v sisteme Arena* [Simulation of economic processes in Arena system], Rostov-on-Don, Rostov State University of Economics, 2012, 128 p.
7. **Cosgrove W. J.** *Simplifying PERT network simulation with Arena, California Journal of Operations Management*, 2008, vol. 6, no. 1, pp. 61—68.
8. **Isaev D. V.** *Analiz programm razvitiya sistem upravleniya s primeneniem programmogo obespecheniya Arena* [Analysis of management systems development programs using Arena software], *Audit and Financial Analysis*, 2015, no. 6, pp. 442—447.
9. **Buntsev I. A.** *Sozdanie i realizatsiya imitatsionnykh modelei v programnoi srede AnyLogic* [Development and implementation of simulation models in AnyLogic program environment], Moscow, Goryachaya Liniya — Telekom, 2015, 153 p.
10. **Emelyanov A. A., Emelyanova N. Z.** *Imitatsionnoe modelirovanie i kompyuternyi analiz ekonomicheskikh protsessov* [Simulation modeling and computer analysis of economic processes], Smolensk, Universum, 2013, 266 p.
11. **Brazhnik A. N.** *Imitatsionnoe modelirovanie: vozmozhnosti GPSS WORLD* [Simulation modeling: capability of GPSS WORLD], Saint Petersburg, Renome, 2006, 439 p.

Критериальное моделирование в виртуализации квалификационных испытаний

Рассмотрены методы виртуализации квалификационных испытаний на основе применения критериальных моделей. Эффективность указанных методов продемонстрирована на примерах моделирования химической стабильности жидкости в гидростатических условиях в целях определения параметров поверочного эксперимента, процедуры оптимизации состава жидкофазной смеси по критерию достижения наилучших характеристик, а также использования результатов прогнозирования фрикционного износа для формирования соответствующих квалификационных норм.

Ключевые слова: квалификационные испытания, полунатурный эксперимент, виртуализация, физическая модель, критериальная модель, квалификационные нормы, химическая трансформация, уравнение Аррениуса, линейная регрессия

Главной особенностью квалификационных испытаний является то, что их основным методом является лабораторное исследование объекта, который, в принципе, ничем не отличается или не будет отличаться от эксплуатируемого [1]. Лабораторную установку при этом можно интерпретировать как реализацию некоторой физической модели системы, включающей исследуемый объект. В этом смысле квалификационные испытания являются частным случаем более широкого понятия полунатурных экспериментов. В рамках этих экспериментов не накладываются ограничения ни на характер модели, которая может быть как физической, так и математической, ни на характер объекта исследования, который может быть как реальным, так и виртуальным. В частности, тестирование программного обеспечения также может осуществляться в рамках полунатурного эксперимента. А поскольку основное свойство исследуемого объекта — неотличимость от эксплуатируемого имеет место и здесь, указанное тестирование также можно интерпретировать как своеобразные квалификационные испытания.

Таким образом, адекватные математические модели системы и объекта можно интерпретировать как виртуальную реальность, а численный эксперимент с их использованием — как виртуальную квалификацию. При этом натурные эксперименты, проводимые в рамках обычных квалификационных испытаний, могут выступать как средство идентификации создаваемых математических моделей. В частности, при наличии физической модели исследуемого процесса в виде лабораторной установки возможной базой для виртуализации являются критериальные модели, базирующиеся на использовании методов теории подобия [2, 3]. Эти методы традиционно применяют в исследованиях физических процессов в ситуации недостаточной проработанности теории последних. Моделирование физических процессов сводится в этом случае к построению их безразмерных инвариантов — критериев подобия, а также соотношений между ними. Как правило, указанные инварианты формируют либо

путем нормализации точных уравнений физических процессов, либо по результатам априорного анализа параметров, влияющих на процесс, с использованием методов теории размерности.

В дальнейшем построенные критериальные модели могут быть использованы для экстраполяции лабораторных результатов на штатные условия эксплуатации объекта, а обратный пересчет, осуществляемый с их помощью, позволяет сформировать условия лабораторного эксперимента, адекватные прогнозируемым эксплуатационным условиям. Отсюда вытекают основные цели виртуализации квалификационных испытаний:

- прогнозирование эксплуатационных характеристик исследуемого объекта по результатам квалификационных испытаний;
- обратный пересчет эксплуатационных условий в адекватные параметры квалификационных испытаний, а эксплуатационных требований — в соответствующие квалификационные нормы.

Моделирование химической стабильности

Цель проведенного моделирования — определение параметров поверочного эксперимента по оценке химической стабильности жидкости в гидростатическом случае [4]. Актуальность сформулированной цели была обусловлена тем, что эксперимент проводили при повышенной температуре для сокращения срока испытаний. Поэтому в число оцениваемых параметров эксперимента была включена его длительность. Таким образом, в ходе проведенного рассмотрения решали обратную задачу виртуализации.

Моделирование осуществляли путем построения следующих критериев подобия:

- геометрического подобия экспериментального и натурального сосудов вида $S/V^{2/3}$, где S , V — поверхность и объем сосуда;
- термодинамического подобия температур и давлений вида $\frac{P}{\rho T}$, где P , T — гидростатическое

давление и температура жидкости, а c , ρ — ее удельная теплоемкость и плотность;

- кинетического подобия протекания химической трансформации в натуральных и экспериментальных условиях вида $\ln \frac{k_0 t}{\mu \rho V \cdot dgr} - \frac{E_0}{RT}$, где k_0 , t , E_0 , dgr — скорость, длительность, энергия активации и достигнутая степень трансформации соответственно, а μ , T — молярная масса и температура жидкости.

Данные критерии определяют параметры квалификационных испытаний, обеспечивающие подобие реальным условиям химической трансформации жидкости, а также длительность этих испытаний. В частности, они позволили рассчитать значение и условия достижения повышенной температуры в испытаниях, обеспечивающей подобие по термодинамическому критерию, а также оценить на основе кинетического критерия длительность этих испытаний, адекватную гарантийному сроку.

Оптимизация состава смеси

Факторный эксперимент в настоящее время используется повсеместно, в том числе в области химической технологии. Наряду с экспериментальным и измерительным оборудованием его существенными составными частями являются методики планирования эксперимента и обработки его результатов. Первая методика подробно рассматривается в специальной литературе [5]. Поэтому здесь мы в основном будем рассматривать методику обработки экспериментальных результатов.

Использование факторной модели предполагает линейно-регрессионную аппроксимацию результатов квалификационного эксперимента. После проведения идентификации модели по результатам калибровочных экспериментов она может быть использована для прогнозирования результатов и планирования контрольных экспериментов. Примером применения факторной модели является оптимизация состава смеси при заданных ограничениях на ее отдельные характеристики. Оптимизация состава смеси осуществлялась с использованием численных методов, работоспособность которых была продемонстрирована в ходе проведенного рассмотрения.

Характеристики смеси определялись долями входящих в нее компонентов, каждый из которых также представляет собой смесь некоторых базовых составляющих. Поскольку в конечном итоге свойства смеси определяются процентным содержанием этих составляющих, целью исследования являлся выбор состава, обеспечивающего оптимальность отдельных характеристик смеси при заданных ограничениях на прочие характеристики.

Предположим, что процентные доли базовых составляющих в составе смеси равны x_1 , x_2 , x_3 , а величины z_1 , z_2 , z_3 представляют собой доли смеси-

ваемых компонентов. Эти доли приведены к допустимым диапазонам содержания составляющих в составе компонентов и связаны с процентным содержанием x_1 , x_2 , x_3 базовых составляющих в составе смеси соотношениями вида:

$$\min_1 + z_1(\max_1 - \min_1) = x_1 \in (\min_1, \max_1);$$

$$\min_2 + z_2(\max_2 - \min_2) = x_2 \in (\min_2, \max_2);$$

$$\min_3 + z_3(\max_3 - \min_3) = x_3 \in (\min_3, \max_3).$$

Пусть результаты измерений характеристик смеси составляют y_1, y_2, y_3 . Примем в соответствии с [1], что влияние процентного состава на каждую характеристику описывается регрессионным соотношением вида:

$$y_i = K_i X = k_{i1}x_1 + k_{i2}x_2 + k_{i3}x_3 + k_{i4}x_1x_2 + k_{i5}x_1x_3 + k_{i6}x_2x_3;$$

$$X = (X_1, X_2, X_3, X_4, X_5, X_6) = (x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3);$$

$$K_i = (k_{i1}, k_{i2}, k_{i3}, k_{i4}, k_{i5}, k_{i6}),$$

$$i = 1, \dots, 3.$$

Расчет вектора факторов X , входящих в состав линейной регрессии, осуществляется по z_1, z_2, z_3 в соответствии со схемой, приведенной на рис. 1. Наряду с линейными факторами регрессии включает их произведения, отображающие взаимное влияние факторов. Коэффициенты линейной регрессии рассчитывают методом неопределенных коэффициентов. С этой целью факторы $X_1—X_6$, сформированные для различных комбинаций z_1, z_2, z_3 и транспонированные, трансформируются в матрицу в соответствии с рис. 2. Эта матрица представляет собой матрицу систем линейных уравнений относительно коэффициентов регрессии. Правые части этих систем образуются векторами y_1, y_2 и y_3 , а их решениями являются наборы коэффициентов

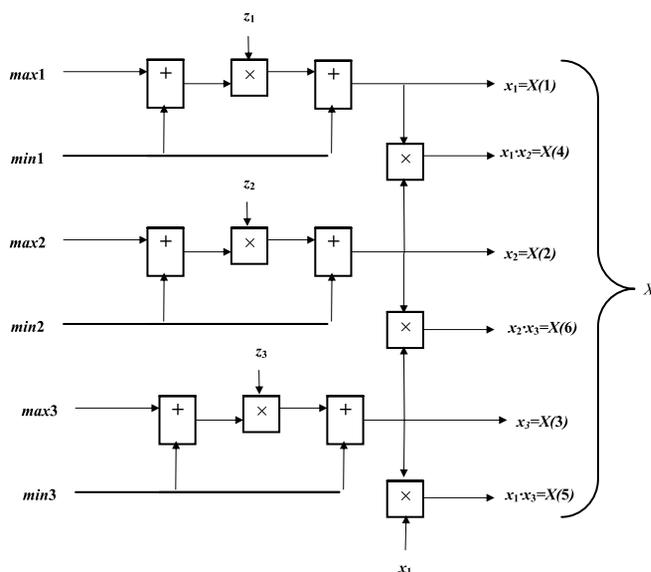


Рис. 1. Расчет регрессионных переменных $R(z_1, z_2, z_3)$

K_1 , K_2 , и K_3 , используемые в дальнейшем для пересчета процентного состава в характеристики смеси.

Задача оптимизации состава, таким образом, сводится к оптимальному выбору приведенных долей z_1 , z_2 , z_3 . В качестве критерия оптимальности в дальнейшем принимается максимум y_1 при условии, что y_2 и y_3 не превышают заданных значений y_{2lim} и y_{3lim} . Реализация соответствующей процедуры условной максимизации приведена на рис. 3. Выбранные значения z_1 , z_2 , z_3 трансформируются в вектор X факторов в модуле $R(z_1, z_2, z_3)$, отображенном на рис. 1. Перемножая этот вектор с векторами коэффициентов K_1 , K_2 и K_3 , формируем текущие характеристики y_1 , y_2 , y_3 . В качестве исходной точки условного максимума принята $(z_1, z_2, z_3) = (1, 0, 0)$. Обновление координат этой точки происходило в ходе перебора z_1 , z_2 , z_3 при выполнении следующего условия С:

$$\begin{aligned} y_1 &> y_{1max}; \\ y_2 &< y_{2lim}; \\ y_3 &< y_{3lim}; \end{aligned}$$

где y_{1max} — текущая оценка максимума y_1 .

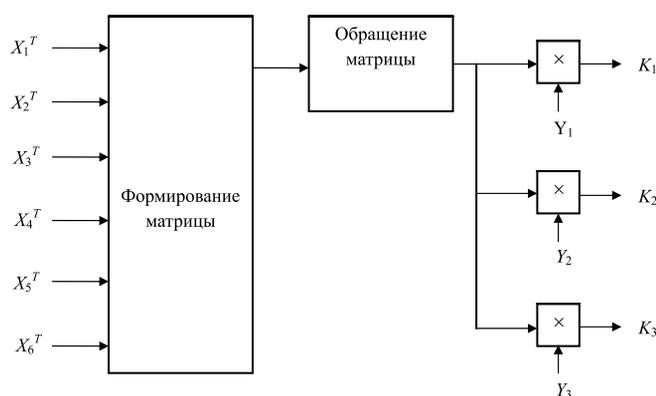


Рис. 2. Формирование регрессионных коэффициентов

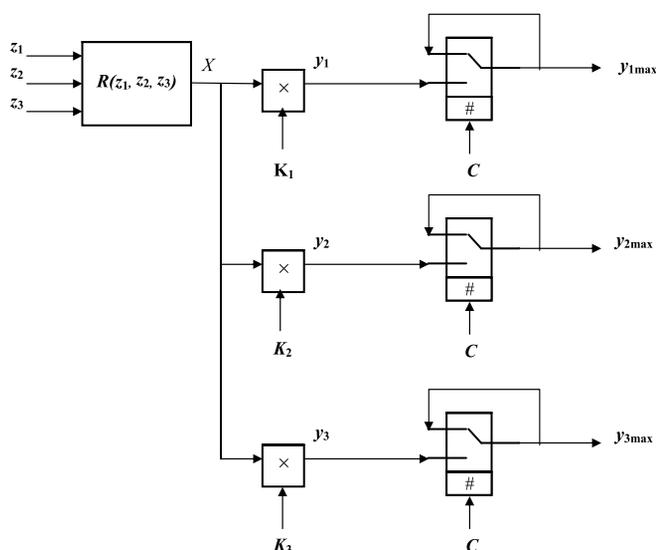


Рис. 3. Оптимизация состава смеси

В рамках реализуемой процедуры рассматривались два метода перебора z_1 , z_2 , z_3 . Первый из них — метод статистического перебора — представляет собой реализацию метода Монте-Карло и отображен на рис. 4. Переменная z_1 представляет собой случайную величину, равномерно распределенную в диапазоне $(0, 1)$, а переменная z_2 — случайную величину, равномерно распределенную в диапазоне $(0, 1 - z_1)$. Переменная z_3 определяется из условия полноты состава смеси $z_1 + z_2 + z_3 = 1$. В случае прямого перебора, отображенном на рис. 5, целочис-

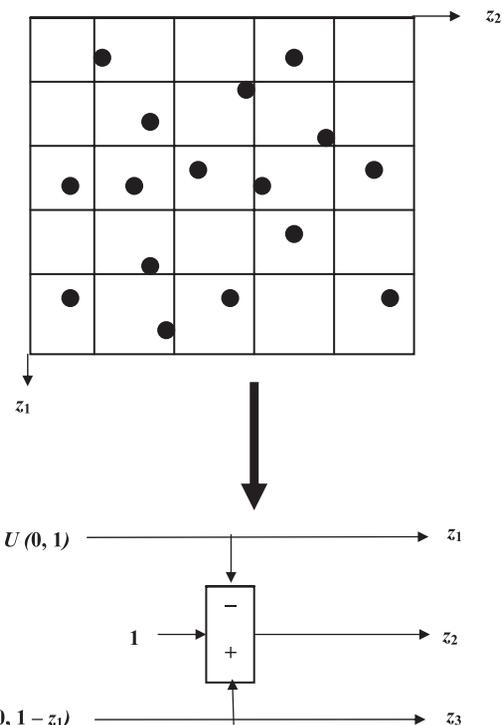


Рис. 4. Статистический перебор

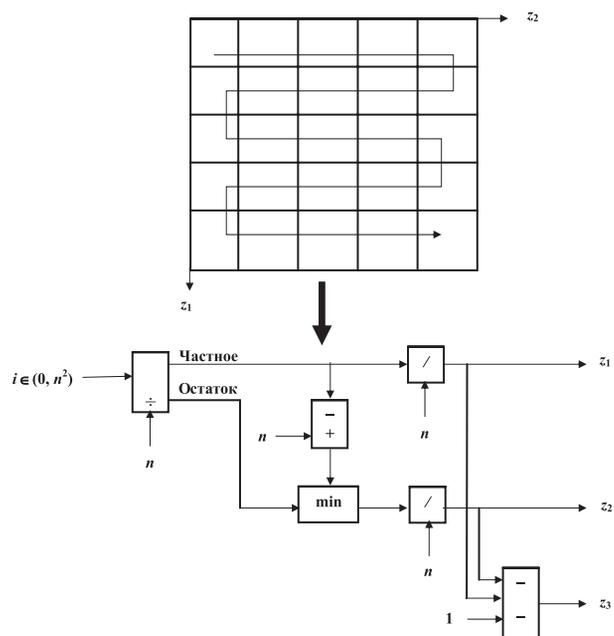


Рис. 5. Прямой перебор

ленная переменная I изменяется в диапазоне $(1, n^2)$. При его делении нацело на n получаем частное и остаток. Частное пересчитывается в z_1 путем деления на n , а остаток сравнивается с величиной n минус частное. Минимальное из этих значений пересчитывается в z_2 путем деления на n . Значение z_3 , как и ранее, находим из условия полноты. В ходе численного расчета, проведенного в соответствии с построенной процедурой, значение n принималось равным 1000. Этому соответствует число вариантов прямого перебора, равное 10^6 . Аналогичное число вариантов было использовано при статистическом переборе. Данный расчет показал, что предложенные процедуры оптимизации состава смеси являются приблизительно равноценными и представляют собой работоспособные средства оптимизации состава смеси в предложенных условиях.

Критериальная модель фрикционного износа

Еще одним примером критериальной модели, используемой для прогнозирования характеристик исследуемого объекта, является модель фрикционного износа плунжерной пары в аксиально-поршневом насосе. На основе анализа основных конструктивных особенностей данного насоса и с учетом противоизносных свойств смазывающего топлива в качестве критерия подобия между лабораторной установкой и исследуемым объектом был выбран критерий объемного износа следующего вида [6]:

$$V_w = \frac{\rho v \Delta V}{\mu N t}. \quad (1)$$

Здесь ρ , ΔV — плотность и дефект объема изнашиваемого материала; μ , N — коэффициент трения и нормальная нагрузка во фрикционной паре; t — длительность износа. Физический смысл данного критерия — отношение импульса увлекаемого износного материала к импульсу силы трения. Он используется для решения следующих задач:

- оптимальный выбор параметров критериальной модели путем минимизации вариации критерия износа;
- формирование уравнения связи между значением износа и параметрами в составе критерия, в частности параметром времени;
- прогнозирование значения эксплуатационного износа при заданных характеристиках пары трения и смазывающего топлива, а также заданной длительности износа;
- расчет параметров квалификационных испытаний и квалификационных норм.
- В работе [6] исследовали схему износа шарообразного измерительного элемента, приведенную на рис. 6. Его объемный износ ΔV соответствует объему шарового сегмента AEB (CFD), равного

$$\Delta V = V_1 - V_2, \quad (2)$$

где V_1 — объем плоского шарового слоя $AEBCFD$, а V_2 — объем усеченного конуса $ABCD$. Рассчитанный объем шарового слоя составил

$$V_1 = \pi R_0^2 (x_2 - x_1) + \frac{\pi}{3} (x_2^3 - x_1^3),$$

$$\text{где } x_1 = \sqrt{R_0^2 - \left(\frac{D}{2}\right)^2}, \quad x_2 = \sqrt{R_0^2 - \left(\frac{d}{2}\right)^2}.$$

Расчетное значение глубины износа ΔL составляет

$$\Delta L = \frac{D_0}{2} \left(1 - \sqrt{1 - \left(\frac{g}{D_0}\right)^2} \right), \quad (3)$$

где ширина полосы износа g соответствует хорде CD и находится из прямоугольного треугольника CDG :

$$g = \frac{D - d}{2 \cos \alpha}. \quad (4)$$

Здесь α — угол наклона оси измерительного шарика ко второму элементу фрикционной пары — конической или цилиндрической шайбе в точке касания.

Критериальная модель износа была реализована в виде рабочего файла в программной среде **Microsoft Excel 7**. Этот файл дополнялся вспомогательным файлом надстройки, который представлял собой библиотеку функций **Visual Basic**, разработанных для расчета показателей износа и используемых для обработки результатов калибровочного эксперимента и прогнозирования характеристик износа моделируемого прототипа в различных условиях. В число данных функций входят:

- функция dV , предназначенная для расчета объемного износа по измерениям максимального и минимального диаметров полосы износа в соответствии с (2);

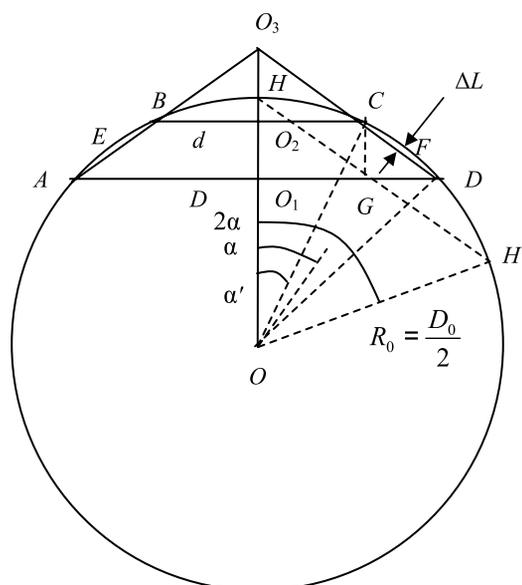


Рис. 6. Износ измерительного шарика

- функция ΔL , предназначенная для расчета глубины износа в соответствии с (3);
- функция g , предназначенная для расчета ширины полосы износа в соответствии с (4);
- функция $So = \frac{\eta v}{E\Delta L(1 - \beta A + \gamma S)}$, предназначенная для расчета числа Зоммерфельда в зависимости от динамической вязкости η , скорости скольжения v , кислотности A и процентного содержания серы S ;
- функция $fr = \mu_0 \exp(-\delta So(\beta, \gamma))$, используемая для расчета коэффициента граничного трения μ по коэффициенту сухого трения μ_0 и числу Зоммерфельда So ;
- функция $vsc = \eta_0 \left(\frac{T}{T_0}\right)^d e^{b\left(\frac{1}{T} - \frac{1}{T_0}\right)}$, используемая для расчета температурной зависимости динамической вязкости и описываемая ниже;
- функция $Vw = \frac{\rho v \Delta V}{\mu N t}$, предназначенная для расчета критерия подобия объемного износа;
- функция $\Delta V_c = \frac{\mu N t}{\rho v} \overline{Vw}$, предназначенная для прогнозирования объемного износа по оптимальному значению критерия подобия \overline{Vw} .

Температурную зависимость динамической вязкости рассчитывали с использованием функции $vsc(T)$, построенной на базе традиционной аппроксимации Андраде $\eta = ce^{\frac{b}{T}}$, где c, b считаются постоянными значениями характеристических вязкости и температуры [8]. Предполагая, что известна стандартная вязкость жидкости, измеренная при постоянной температуре $T_0 = 20^\circ\text{C} = 293\text{ K}$,

получаем $\eta_0 = ce^{\frac{b}{T_0}}$ или $c = \eta_0 e^{-\frac{b}{T_0}}$. Подставляя величину c в аппроксимацию Андраде, получаем выражение текущей вязкости через стандартную

$\eta = \eta_0 e^{b\left(\frac{1}{T} - \frac{1}{T_0}\right)}$. Учитывая возможную зависимость характеристической вязкости от температуры, воспользуемся для построения модифицированной аппроксимации температурной зависимости более общим, чем аппроксимация Андраде, выражением

вида $\eta = \eta_0 \left(\frac{T}{T_0}\right)^d e^{b\left(\frac{1}{T} - \frac{1}{T_0}\right)}$. Будем в дальнейшем

считать, что данное выражение аппроксимирует заданную табличную зависимость $\eta_i = \eta(T_i)$, $i = 1, \dots, n$, где n — число опорных точек зависимости. При логарифмировании выражения в этих опорных точках, оно преобразуется к линейной функции, параметры которой b и d рассчитывают методом наи-

меньших квадратов [9]. Будучи подставлены в полученное выше выражение, они полностью определяют искомую температурную зависимость.

Численный расчет в соответствии с предложенной методикой был выполнен для табличной зависимости вязкости от температуры для керосина [10]. Значения параметров составили для нее $b = 2548\text{ K}$ и $d = 3,156$, а результаты аппроксимации отображены на рис. 7 (см. третью сторону обложки). Они показывают, что обобщенная аппроксимация обеспечивает лучшее приближение для табличной зависимости, чем аппроксимация по Андраде. Поэтому ее предполагается использовать в дальнейшем в составе модели износа.

Моделирование износа включает два этапа:

- параметризация модели по результатам калибровочного эксперимента;
- прогнозирование износа при заданных условиях и длительности.

Схема оптимальной параметризации модели отображена на рис. 8. В ходе калибровочного эксперимента на лабораторной установке осуществляли замеры минимального d и максимального D диаметров полосы износа измерительного элемента [7]. Для их обработки в программу вводили ряд параметров. Часть из них являлась параметрами квалификационных испытаний, к числу которых относились нагрузка используемой установки и время износа, характеристики измерительного элемента, характеристики смазывающего топлива, включая температуру, определяющую динамическую вязкость, кислотность и процентное содержание серы, кинематические параметры установки, определяющие скорость скольжения и др. Другая часть представляла собой параметры температурной зависимости для вязкости смазывающего топлива, описываемой ниже.

Первым этапом параметризации модели являлось формирование статистики критерия Vw по полученным измерениям и варьируемым значениям калибровочных параметров. Применяли следующую схему этого формирования. Параметры выбирали случайно в диапазонах $\delta \in (\delta_{mn}, \delta_{mx})$, $\beta \in (\beta_{mn}, \beta_{mx})$, $\gamma \in (\gamma_{mn}, \gamma_{mx})$. На этом этапе для каждого измерения максимального и минимального диаметров полосы износа, а также варьируемых параметров β, γ (коэффициенты при кислотности и содержании серы в функции So) и варьируемого параметра δ (коэффициента при So в функции fr) рассчитывают следующие величины:

- объемный износ ΔV с использованием функции $dV(D_0, D, d)$;
- глубина ΔL и полоса g с использованием функций ΔL и g ;
- число Зоммерфельда So для расчетной глубины износа ΔL с использованием функции So ;
- коэффициент граничного трения μ с использованием функции fr ;

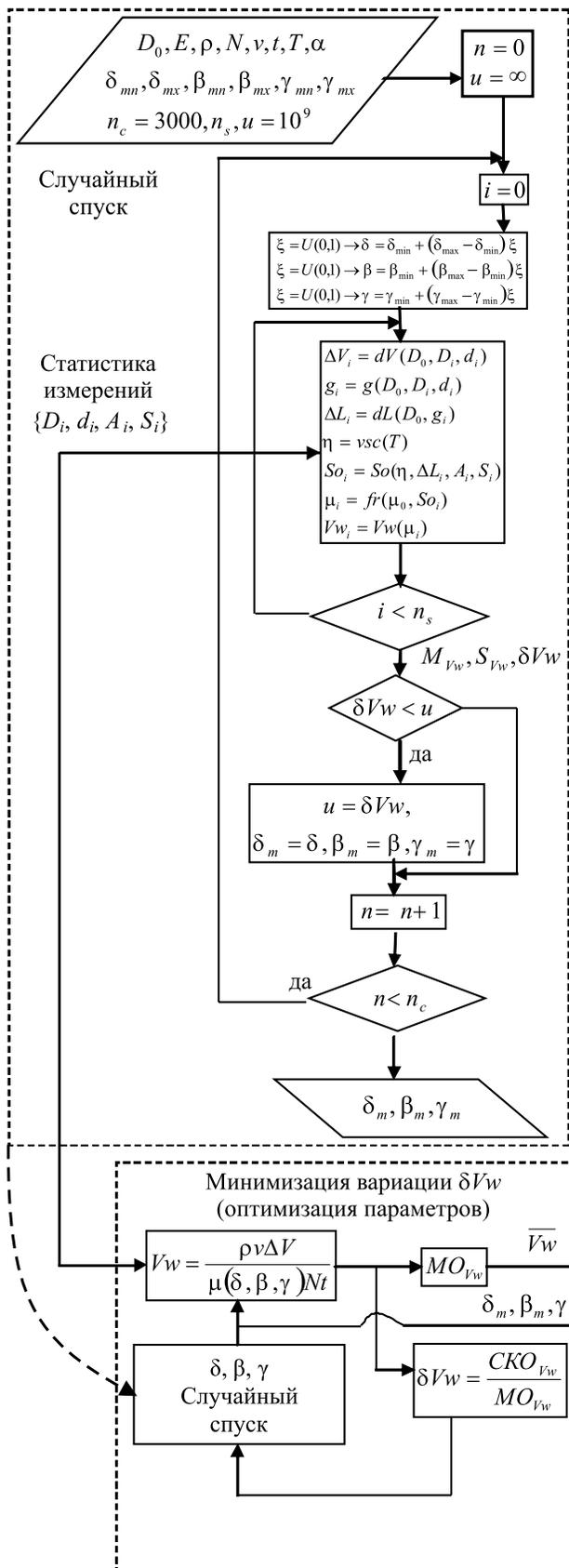


Рис. 8. Оптимальная параметризация

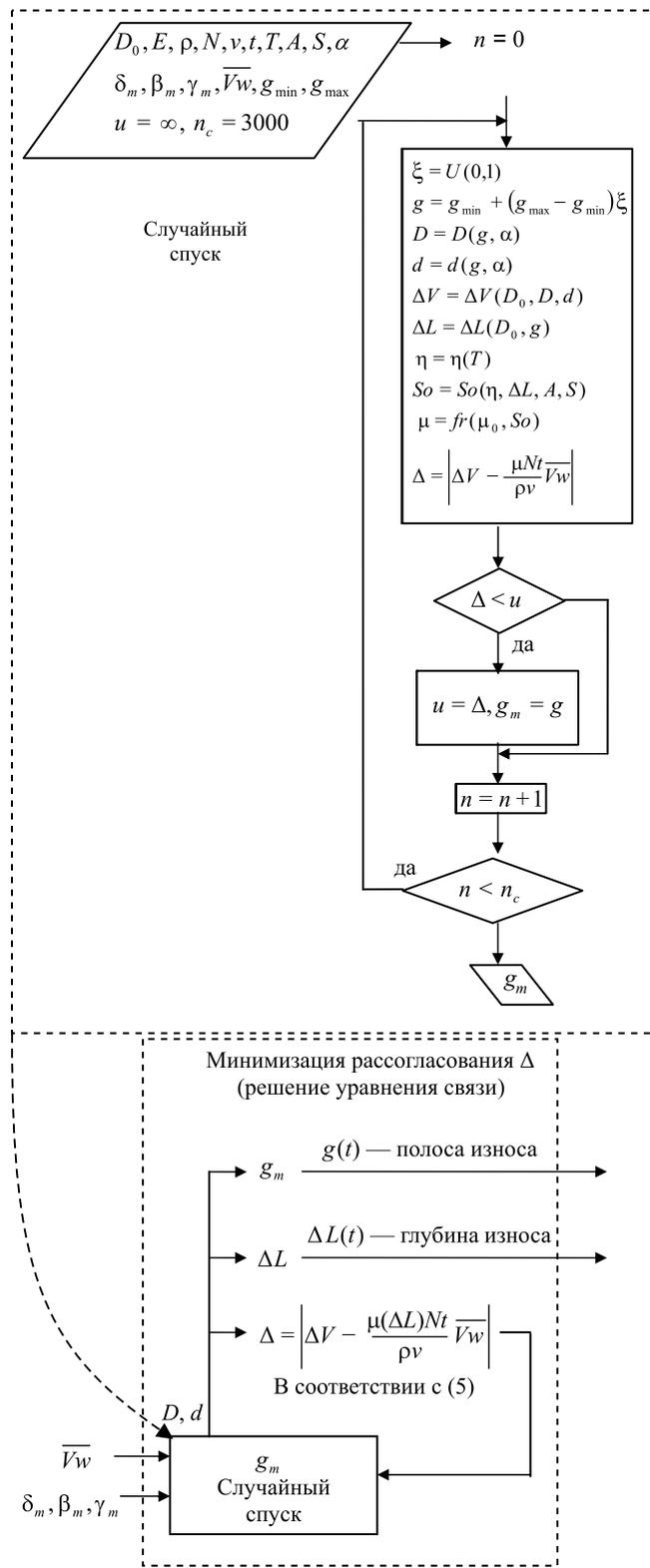


Рис. 9. Прогнозирование износа

- и, наконец, критерий объемного износа V_w с использованием функции V_w .

В совокупности расчетные значения V_w образуют статистику, используемую для расчета среднего значения M_{V_w} и СКО S_{V_w} критерия, а также для расчета вариации критерия $\delta V_w = M_{V_w}/S_{V_w}$ для выбранных значений параметров δ, β, γ . Оптимизация этих параметров осуществляется из следующих соображений. Поскольку критерий подобия можно рассматривать как инвариант критериальной модели, его вариация δV_w при должном выборе параметров должна быть минимальной. Таким образом, оптимальные параметры $\delta_m, \beta_m, \gamma_m$ представляют собой точку минимума вариации в пространстве этих параметров.

В ходе поиска точки минимума вариация δV_w , рассчитанная при текущих значениях параметров δ, β, γ , сравнивается с текущим значением минимума u . В случае если выполняется неравенство $\delta V_w < u$, корректируются текущие значения минимума $u = \delta V_w$ и точки минимума $(\delta_m, \beta_m, \gamma_m) = (\delta, \beta, \gamma)$. В качестве оценки точки минимума принимали значение, достигнутое к концу цикла. К числу параметров относилась также величина математического ожидания критерия подобия $\bar{V}_w = M_{V_w}$, соответствующая оптимальным параметрам $\delta_m, \beta_m, \gamma_m$.

После завершения обработки результатов калибровочного эксперимента возможен переход к прогнозированию износа в соответствии со схемой, приведенной на рис. 9. Этому предшествовал ввод в программу скорректированных значений параметров, а также значений кислотности и процентного содержания серы. Прогнозирование износа осуществляется на основе следующих соотношений:

$$\begin{cases} \Delta V_c = \frac{\mu N t}{\rho v} \bar{V}_w, \\ \mu = \mu_0 \exp(-\delta S_0(\beta, \gamma)), \\ S_0 = \frac{\eta v}{E \Delta L (1 - \beta A + \gamma S)}. \end{cases} \quad (5)$$

В указанные соотношения подставляли параметры модели, выбранные на этапе параметризации. После этой подстановки первое соотношение (5) интерпретируется как уравнение связи между расчетным износом ΔV_c и прочими параметрами критерия. Это уравнение решали путем минимизации модуля рассогласования Δ между правой и левой частями уравнения для износа, результатом которой являются зависимости глубины износа $\Delta L(t)$ и полосы износа $g(t)$ от времени, которые интерпретируются как прогнозы износа.

Объемный износ ΔV в левой части уравнения (5), а также глубина износа ΔL , входящая в выражение для числа Зоммерфельда, являются в соответствии с (2), (3) функциями минимального d и максимального D диаметров полосы износа. При расчете этих диаметров в соответствии с рис. 6 считается, что перпендикуляр к плоскости шайбы в точке касания

шарика F совпадает с биссектрисой угла COD и $\arcsin \frac{D}{D_0} = \angle O_3 OD = \alpha + (\alpha - \alpha') = 2\alpha - \alpha'$.

Учитывая, что $\arcsin \frac{d}{D_0} = \alpha'$, получаем $\arcsin \frac{d}{D_0} + \arcsin \frac{D}{D_0} = 2\alpha$. Отсюда находим с учетом (3)

уравнения для максимального D и минимального d диаметров полосы износа:

$$\begin{cases} \arcsin \frac{D}{D_0} + \arcsin \frac{d}{D_0} = 2\alpha, \\ g = \frac{D-d}{2 \cos \alpha}, \end{cases}$$

решения которых имеют следующий вид:

$$\begin{cases} D = D(g, \alpha) = \sin \alpha \sqrt{D_0^2 - g^2} + g \cos \alpha, \\ d = d(g, \alpha) = \sin \alpha \sqrt{D_0^2 - g^2} - g \cos \alpha. \end{cases}$$

Эти решения определяют объемный износ ΔV в соответствии с (2), глубину износа ΔL в соответствии с (3), и текущее рассогласование правой и левой частей первого уравнения (5). Для обеспечения минимальности этого рассогласования при заданных характеристиках пары трения и заданной длительности износа величина g в диапазоне $g \in (0, g_{\max})$ подбирается, как и ранее, методом случайного спуска. При этом согласно рис. 6 максимальное значение диапазона $g_{\max} = D_0 \sin \alpha$ соответствует длине хорды NN' , противолежащей центральному углу 2α . Как и ранее, условию $\Delta < u$ соответствует обновление текущего минимума $u = \Delta$ и текущей точки минимума $g_m = g$, а результатом минимизации считается достигнутое значение g_m .

Разработанная модель может быть использована для формирования квалификационных норм на основе соответствующих эксплуатационных требований. Примером такого требования является, например, ограничение на предельную глубину износа плунжера, которая по некоторым оценкам не должна превышать 150 мкм при длительности износа 1000 ч. В качестве примера реализации разработанной процедуры прогнозирования были рассмотрены прогнозы износа плунжера и измерительного элемента при предполагаемых характеристиках квалификационных испытаний. В перспективе это дает возможность оценить точность прогнозирования полосы износа путем сравнения с прямыми измерениями на лабораторной установке.

Для моделирования износа плунжера использовали следующие его характеристики:

- диаметр головки плунжера — 9 мм;
- удельная нагрузка — 60 атм;
- площадь контакта — 0,4 мм²;
- скорость скольжения — 0,1 м/с;
- интегральная нагрузка — 2,4 Н.

Таблица 1

Перевод твердости

Роквелл		Бринелль		Виккерс HV
HRC	HRA	Диаметр от- печатка, мм	НВ при нагрузке 300 кгс (3000 Н)	
72	89,0	2,20	782	1220
70	86,5	—	—	1076
69	86,0	2,25	744	1004
68	85,5	—	—	942
67	85,0	2,30	713	894
66	84,5	—	—	854
65	84,0	2,35	683	820
65	83,5	—	—	789
63	83,0	2,40	652	763
62	82,5	—	—	739
61	81,5	2,45	627	715
60	81,0	—	—	695
59	80,5	2,50	600	675
58	80,0	2,55	—	655
57	79,5	—	578	636
56	79,0	2,60	—	617
55	78,5	—	555	598
54	78,0	2,65	—	580
53	77,5	—	532	562
52	77,0	2,70	512	545
51	76,5	2,75	495	528
50	76,0	—	—	513
49	75,5	2,80	477	498
48	74,5	2,85	460	485
47	74,0	2,89	448	471
46	73,5	2,92	437	458
45	73,0	2,96	426	446
44	72,5	3,00	415	435
42	71,5	3,08	393	413
40	70,5	3,16	372	393
38	—	3,25	352	373
36	—	3,34	332	353

Твердость материала головки плунжера и измерительного шарика по Роквеллу принимали равной 65 HRC и переводили в твердость по Бринеллю, используемую в расчетах, в соответствии с табл. 1 (<http://promotreid.ru/promotreidteinfo/pro->

motreidtv/). Были также использованы следующие характеристики топлива:

- марка топлива ТС-1, ГОСТ 10227—86;
- плотность 789,3 кг/м³;
- кислотность $A = 0,12$ мг КОН на 100 см³ топлива;
- массовое содержание серы $S = 0,131$ %.

Считается, что при равных удельных нагрузках интегральная нагрузка плунжера пропорциональна квадрату кривизны головки, т. е. связана с эквивалентной нагрузкой лабораторной установки соот-

ношением $\frac{N_p}{D_p^2} = \frac{N}{D_0^2}$, где D_0, D_p — диаметры изме-

рительного шарика и головки плунжера. Используя это соотношение, можно оценить нагрузку лабораторной установки, соответствующую нагрузке плунжера. В частности, диаметру шарика 0,0127 м соответствует эквивалентная нагрузка 5 Н, диаметру шарика 0,02 м — эквивалентная нагрузка 12 Н, диаметру шарика 0,0254 м — эквивалентная нагрузка 20 Н. Прочие параметры установки приведены ниже.

Время износа, ч	1...2000
Скорость скольжения, м/с	0,1
Твердость по Роквеллу, HRC	65
Температура эксперимента, °С	50
Наклон оси измерительного шарика, °	10
Диаметр измерительного шарика, м	0,0127, 0,02, 0,0254
Диаметр измерительной шайбы, м	0,04

Результаты прогнозирования износа головки плунжера и измерительных шариков лабораторной установки при эквивалентных интегральных нагрузках приведены в табл. 2. Их анализ показывает, что значение износа при его длительности 1000 ч не превышает 150 мкм, т. е. приблизительно соответствуют принятой норме на глубину износа. Непосредственно измеряемыми в квалификационном эксперименте являются значения полосы износа, соответствующие времени эксперимента 1 ч и выделенные жирным шрифтом. Их и предлагается принять в качестве квалификационных норм.

Таблица 2

Прогноз износа

Длительность износа, ч	Головка плунжера 9 мм, нагрузка 2,4 Н		Измерительный шарик 25,4 мм, нагрузка 20 Н		Измерительный шарик 20 мм, нагрузка 12 Н		Измерительный шарик 12,7 мм, нагрузка 5 Н	
	Полоса износа, мм	Глубина износа, мкм	Полоса износа, мм	Глубина износа, мкм	Полоса износа, мм	Глубина износа, мкм	Полоса износа, мм	Глубина износа, мкм
1	0,103787708	0,418934067	0,481958376	3,201326953	0,334684965	1,960514507	0,165700963	0,756746846
10	0,211691787	1,743216526	0,88991219	10,91918178	0,638557862	7,139300369	0,347980739	3,338370932
100	0,395137643	6,077614133	1,606386496	35,62772519	1,157857863	23,49976628	0,631988932	11,02075143
500	0,629624082	15,45373496	2,570376129	91,50051128	1,849088861	60,08750385	1,021586202	28,8534614
1000	0,793290775	24,56706062	3,23898011	145,7314957	2,311904594	94,15637958	1,281182564	45,46408208
2000	1,000089138	39,13405145	4,086386861	233,0929288	2,934410055	152,3122735	1,605194581	71,5747638

Выводы

1. Виртуализация квалификационных испытаний базируется на использовании математического моделирования процесса испытаний, в частности, с применением методов теории подобия, и решает следующие основные задачи:

- экстраполяция результатов квалификационных испытаний на штатные условия эксплуатации;
- обратный пересчет эксплуатационных условий в адекватные параметры квалификационных испытаний;
- формирование квалификационных норм по результатам прогнозирования эксплуатационных характеристик.

2. С использованием алгебраических методов теории подобия были синтезированы критерии химической стабильности жидкости в гидростатических условиях — критерии геометрического, термодинамического и кинетического подобия. Эти критерии, с одной стороны, позволили установить параметры квалификационных испытаний, адекватные условиям эксплуатации. С другой стороны, на их основе анализируется влияние изменений условий проведения эксперимента на корректность прогнозирования эксплуатационных характеристик.

3. Предложена методика оптимизации состава жидкофазной смеси с использованием результатов факторных экспериментов для предварительной калибровки линейно-регрессионных моделей характеристик смеси. Собственно оптимизация осуществляется расчетным путем с использованием откалброванных моделей характеристик и численным варьированием состава смеси методом Монте-Карло или прямого перебора.

4. При исследовании фрикционного износа плунжерной пары алгебраический подход был использован для построения критерия объемного износа. Оптимальная параметризация данной критериальной модели базировалась на интерпретации критерия износа как инварианта, а оптимальных параметров — как точки минимума вариации кри-

терия. При этом соотношение, связывающее критерий объемного износа и его расчетное значение с минимальной вариацией, интерпретировалось как уравнение связи между значением износа и прочими параметрами критерия, включая время износа.

5. Построенное уравнение связи используется для прогнозирования значения износа в заданных условиях и для различных сроков эксплуатации. Результаты прогнозирования предложено использовать для формирования квалификационных норм износа на основании эксплуатационных ограничений на глубину износа при заданной длительности последнего. С помощью критериальной модели рассчитывалась ширина полосы износа измерительного элемента при заданной длительности эксперимента, которую в перспективе предлагается принять в качестве квалификационной нормы при проведении испытаний.

Список литературы

1. **ГОСТ 16504—81.** Испытания и контроль качества продукции. Основные термины и определения. М.: Издательство стандартов, 1991.
2. **Kline S.** Similitude and approximation theory. Berlin — Heidelberg: Springer, 1986. 229 p.
3. **Barenblatt G.** Scaling, self-similarity and intermediate asymptotics. Cambridge: Cambridge university press, 1996. 383 p.
4. **Моисеев А. А.** Моделирование химической стабильности методом теории подобия // Информационные технологии. Том 21, № 9. 2015. С. 666—669.
5. **Ахназарова С. Л., Кафаров В. В.** Методы оптимизации эксперимента в химической технологии. М.: Высшая школа, 1985. 327 с.
6. **Моисеев А. А.** Критериальная модель квалификационных испытаний на износ // Инженерная физика. 2015. № 12. С. 30.
7. **Никитин И. М., Кондратенко В. В., Баевский Д. Ф.** Сравнительная оценка определения противозносных свойств отечественных реактивных топлив на установке ВОСЛЕ и приборе ПСТ — 2 // Труды ГОСНИИ химмотологии. М.: Перо, 2014. Вып. 56. С. 175.
8. **Reid R., Prausnitz J., Sherwood T.** The properties of gases and liquids. NY: McGraw-Hill, 1977. 592 p.
9. **Линник Ю. В.** Метод наименьших квадратов и основы математико-статистической обработки наблюдений. М.: Физматгиз, 1958. 336 с.
10. **Физические величины:** справочник / под ред. И. С. Григорьева и Е. З. Мейлихова. М.: Энергоатомиздат, 1991. 1232 с.

A. A. Moiseev, PhD, Senior Researcher, slow.coach@yandex.ru, State Research Institute of Chimnotology

Criteria Modeling in Virtualization of Qualification Tests

Considered the methods of qualification tests virtualization on the base of criteria models application. Their effectiveness demonstrated by the example of chemical stability description at hydrostatic conditions, liquid-phase mixture composition optimization and plunger piston friction wear. Chemical stability modeling used for determination of qualification tests conditions. This modeling performed by means of similarity criteria building: geometrical, thermodynamic and kinetic ones. These criteria allowed setting qualification tests conditions which are similar to real conditions of liquid chemical transformation and also qualification test durability. Developed procedure of liquid-phase mixture composition optimization based on mixture characteristics presentation in form of linear regression from mixture combination. This optimization based on mixture components fractions variation. At friction wear modeling were proposed rational methods of tests conditions choice and also optimization and forecasting methods for plunger wear characteristics in these conditions. It was proposed also to use these results for qualification standards formation.

Keywords: qualification tests, semi-natural experiment, virtualization, physical model, criteria model, qualification standards, chemical transformation, Arrhenius equation, linear regression

References

1. **GOST 16504—81.** Ispytaniya i control kachestva produktsii. Osnovnye terminy i opredeleniya. (Testing and control of production quality. Base terms and definitions), Moscow, Publishing House of standards, 1991.
2. **Kline S.** Similitude and approximation theory, Berlin — Heidelberg, Springer, 1986, 229 p.
3. **Barenblatt G.** Scaling, self-similarity and intermediate asymptotic, Cambridge: Cambridge university press, 1996. 383 p.
4. **Moiseev A.** Modelirovaniye chimicheskoy stabilnosti metodom teorii podobiya (Similarity relations in chemical stability analysis), *Informacionnye tehnologii*, vol. 21, no. 9, 2015, pp. 666—669.
5. **Achnazarova S., Kafarov V.** Metody optimizatsii experimenta v chimicheskoy tehnologii (Experiments optimization methods in chemical technology), M., Vysshaya skola, 1985, 327 p.
6. **Moiseev A.**, Kriteriynaya model kvalifikatsionnykh ispytaniy na iznos (Criteria model of wear qualification tests), *Ingenernaya fizika*, 2015, no. 12, p. 30.
7. **Nikitin I., Kondratenko V., Baevskiy D.** Sravnitel'naya otsenka opredeleniya protivoiznosnykh svoystv otechestvennykh reaktivnykh topliv na ustanovke BOCLE i pribore ПСТ — 2 (Comparative estimation of antiwear properties for domestic jet — fuel using BOCLE and ПСТ — 2 plants), *Trudy GOSNII himmotologii* (Proceedings of State institute of himmotology), Moscow, Pero ("Pen"), 2014, issue 56, p. 175.
8. **Reid R., Prausnitz J., Sherwood T.** The properties of gases and liquids, NY, McGraw-Hill, 1977, 592 p.
9. **Linnik Yu.** Method naimenshich kvadratov i osnovy matematiko-statisticheskoy obrabotki nabludeniy (Least square method and fundamentals of observations mathematical-statistical processing), M., publisher "Fizmatgiz", 1958, 336 p.
10. **Fizicheskiye velichiny: spravochnik** (Physical values: reference book), ed. by I. Grigor'ev, E. Meilichov, Moscow, Energoatomizdat, 1991, 1232 p.

УДК 621.039.546

Е. П. Доморацкий, д-р техн. наук, проф., e-mail: Domorackiy.E@mfua.ru
Московский финансово-юридический университет МФЮА,
Финансовый университет при Правительстве Российской Федерации, г. Москва

Алгоритм синтеза геометрических характеристик проекционных изображений трехмерных объектов

Рассмотрен алгоритм синтеза геометрических характеристик одномерных и двумерных проекционных изображений выпуклых трехмерных объектов неправильной формы, аппроксимируемых эллипсоидом общего вида. Алгоритм содержит математические модели последовательности операций (процедур) получения проекционных изображений, их морфологического анализа (получения базовых геометрических признаков) и получения численных значений пространственных геометрических характеристик проекционных изображений.

Ключевые слова: алгоритм, синтез, геометрические характеристики, объект, проекционное изображение, математическая модель, аппроксимация, эллипсоид, базовые признаки

Введение

В настоящее время в различных областях науки, техники и практики широко применяются методы и средства определения геометрических характеристик разного рода объектов, основанных на различных физических принципах. К их числу относятся томографические, оптикоэлектронные, телевизионные, радиографические и другие методы. Все эти методы базируются на получении, обработке, анализе и синтезе различного рода одномерных и двумерных проекционных изображений объекта [1—3].

Постановка задачи (цель работы)

Данная работа относится к области морфологической обработки изображений, и ее цель заключается в синтезе (разработке) математических моделей последовательности операций (процедур) получения и преобразования исходного (синтаксического)

проекционного изображения выпуклого трехмерного объекта неправильной формы в семантическое (смысловое) изображение путем выбора и определения его оптимальных базовых геометрических признаков и получения на этой основе прагматического изображения, т.е. численных значений геометрических характеристик размеров и формы одно- и двумерных проекционных изображений объекта [4, 5].

Описание процедур алгоритма синтеза

Алгоритм синтеза включает следующие операции (этапы).

1. Выбор математической модели, аппроксимирующей выпуклый трехмерный объект неправильной формы, и геометрических характеристик его размеров и формы.

2. Определение и выбор оптимальных базовых геометрических признаков проекционных изображений.

3. Получение математической модели двумерного проекционного изображения на выбранную плоскость наблюдения (ракурс) и геометрических характеристик его размеров и формы.

4. Получение математических моделей зависимостей выбранных базовых признаков проекционного изображения от геометрических размеров и пространственной ориентировки (ракурсов) трехмерного объекта.

5. Определение геометрических характеристик размеров и формы проекционного изображения произвольного ракурса и получение зависимостей их численных значений от выбранных базовых геометрических признаков.

Рассмотрим операции данного алгоритма более подробно.

В данной работе в качестве математической модели, аппроксимирующей геометрические характеристики выпуклого трехмерного объекта неправильной формы (его трехмерное изображение — образ), выбран эллипсоид общего вида. При этом геометрические (габаритные) размеры трехмерного объекта определяются численными значениями взаимно ортогональных осей эллипсоида $2A_1, 2A_2, 2A_3$, а его средний диаметр (D) и фактор формы (K) определяются соответственно из следующих соотношений:

$$D = \sum_{i=1}^3 2A_i/3; K = \frac{\max\{A_1, A_2, A_3\}}{\min\{A_1, A_2, A_3\}}.$$

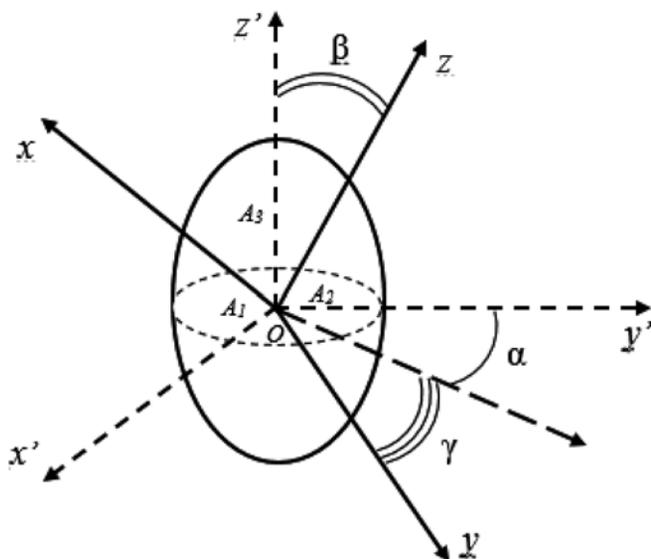


Рис. 1. Схема поворота системы координат относительно центра эллипсоида

Оптимальные базовые признаки проекционных изображений выберем на основе численного анализа их информативности методом максимальной энтропии [6, 7]. К ним относятся: длина одномерной проекции объекта на выбранное направление h_i ; максимальный $l_{\max i}$ и минимальный $l_{\min i}$ линейные габаритные размеры двумерной проекции; площадь двумерной проекции s_i (i — номер проекции). Число проекционных изображений и их ракурсы определяются исходя из требований к точности и достоверности определения пространственных геометрических характеристик (размеров и формы) трехмерного объекта (аппроксимируемого эллипсоидом общего вида) [4, 8].

Получим зависимости выбранных базовых признаков проекционных изображений от линейных размеров $2A_1, 2A_2, 2A_3$ произвольно ориентированного в пространстве эллипсоида общего вида.

Каноническое уравнение эллипсоида общего вида в декартовой системе координат имеет вид

$$(x'^2/A_1^2) + (y'^2/A_2^2) + (z'^2/A_3^2) = 1, \quad (1)$$

где A_1, A_2, A_3 — полуоси эллипсоида; x', y', z' — исходная система координат эллипсоида.

Поворот эллипсоида в пространстве относительно его центра эквивалентен повороту системы координат относительно этого же центра и может быть представлен в виде трех последовательных поворотов вокруг осей координат (рис. 1). В этом случае матрица R_0 , характеризующая поворот, может быть выражена произведением трех матриц, описывающих последовательные повороты вокруг осей z, y и z на соответствующие углы Эйлера α, β, γ [9]:

$$R_0 = R_z(\gamma)R_y(\beta)R_z(\alpha),$$

где

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}; R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix};$$

$$R_z(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Элементы матрицы R_0 позволяют выразить координаты точки в повернутой системе координат $Ox'y'z'$ через координаты точки в неповернутой системе $(Ox'y'z')$. Обратное преобразование определяется матрицей R :

$$R = R_0^{-1} = R_z(-\alpha)R_y(-\beta)R_z(-\gamma).$$

Учитывая, что функция $f(x) = \cos(x)$ является четной, а функция $f(x) = \sin(x)$ является нечетной, получим:

$$R_z(-\alpha) = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}; R_y(-\beta) = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix};$$

$$R_z(-\gamma) = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Таким образом, матрица поворота

$$R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta & \sin\alpha & -\cos\alpha\sin\beta \\ -\sin\alpha\cos\beta & \cos\alpha & \sin\alpha\sin\beta \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & \cos\alpha\cos\beta\sin\gamma + \sin\alpha\sin\gamma & -\cos\alpha\sin\beta \\ -\sin\alpha\cos\beta\cos\gamma - \cos\alpha\sin\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta \\ \sin\beta\cos\gamma & \sin\beta\sin\gamma & \cos\beta \end{bmatrix}.$$

Матрица R задает преобразование:

$$\begin{cases} x' = R_{11}x + R_{12}y + R_{13}z; \\ y' = R_{21}x + R_{22}y + R_{23}z; \\ z' = R_{31}x + R_{32}y + R_{33}z. \end{cases} \quad (2)$$

Подставляя выражение (2) в (1) и группируя члены уравнения (1) относительно произведений переменных x, y, z , получаем уравнение эллипсоида в повернутой системе координат, аналогичное уравнению эллипсоида, повернутого в пространстве:

$$\begin{aligned} & \left[\sum_{i=1}^3 (R_{i1}^2/A_i^2)x + \sum_{i=1}^3 (R_{i1}R_{i2}/A_i^2)y + \right. \\ & \left. + \sum_{i=1}^3 (R_{i1}R_{i3}/A_i^2)z \right] + \left[\sum_{i=1}^3 (R_{i2}R_{i2}/A_i^2)x + \right. \\ & \left. + \sum_{i=1}^3 (R_{i2}^2/A_i^2)y + \sum_{i=1}^3 (R_{i2}R_{i3}/A_i^2)z \right] + \\ & \left[\sum_{i=1}^3 (R_{i3}R_{i3}/A_i^2)x + \sum_{i=1}^3 (R_{i3}R_{i3}/A_i^2)y + \right. \\ & \left. + \sum_{i=1}^3 (R_{i3}^2/A_i^2)z \right] = 1. \end{aligned} \quad (3)$$

Из формулы (3) получим уравнение касательной плоскости к эллипсоиду в произвольной точке (x_1, y_1, z_1) его поверхности и наложим условие перпендикулярности этой плоскости к плоскости xOy . В этом случае уравнение касательной плоскости не будет зависеть от координаты z и примет вид

$$ax + by = 1, \quad (4)$$

где

$$\begin{cases} a = \sum_{i=1}^3 (R_{i1}^2/A_i^2)x_1 + \sum_{i=1}^3 (R_{i1}R_{i2}/A_i^2)y_1 + \\ + \sum_{i=1}^3 (R_{i1}R_{i3}/A_i^2)z_1; \\ b = \sum_{i=1}^3 (R_{i2}R_{i2}/A_i^2)x_1 + \sum_{i=1}^3 (R_{i2}^2/A_i^2)y_1 + \\ + \sum_{i=1}^3 (R_{i2}R_{i3}/A_i^2)z_1; \\ 0 = \sum_{i=1}^3 (R_{i3}R_{i3}/A_i^2)x_1 + \sum_{i=1}^3 (R_{i3}R_{i3}/A_i^2)y_1 + \\ + \sum_{i=1}^3 (R_{i3}^2/A_i^2)z_1. \end{cases} \quad (5)$$

Исключая из системы (5) z_1 и учитывая, что выражение (4) должно выполняться в точке (x_1, y_1, z_1) , подставляем уравнения для a и b в (4) и приводим его к виду

$$\begin{aligned} & \sum_{i=1}^3 A_i^2 R_{i2}^2 x_1^2 - 2 \sum_{i=1}^3 A_i^2 R_{i1} R_{i2} x_1 y_1 + \sum_{i=1}^3 A_i^2 R_{i1}^2 y_1^2 = \\ & = A_1^2 A_2^2 A_3^2 \sum_{i=1}^3 (R_{i3}^2/A_i^2). \end{aligned} \quad (6)$$

Так как точка (x_1, y_1, z_1) была взята произвольно, уравнение (6) в совокупности с третьим уравнением системы (5) однозначно определяет набор точек на поверхности эллипсоида, в которых плоскости, касательные к эллипсоиду, являются перпендикулярными к координатной плоскости xOy . Соответствующий им набор точек $(x_1, y_1, 0)$ определяет контур проекции эллипсоида на координатную плоскость xOy , являющийся эллипсом вида (6). На рис. 2 показано проекционное изображение эллипсоида общего вида на координатную плоскость xOy и его базовые геометрические признаки.

Уравнение (6) для координатной плоскости xOy представим в виде

$$ax^2 - 2bxy + cy^2 = d, \quad (7)$$

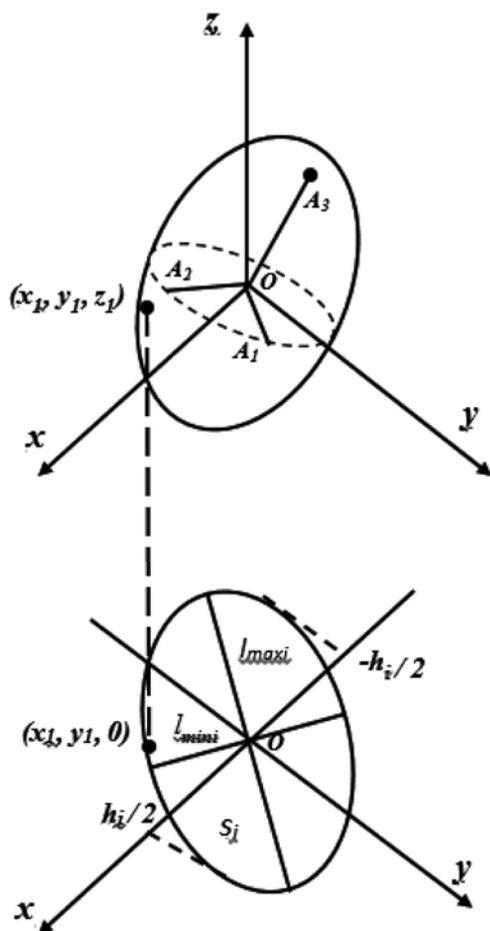


Рис. 2. Проекционное изображение эллипсоида общего вида на координатную плоскость xOy

где

$$\begin{aligned}
 a &= \sum_{i=1}^3 A_i^2 R_{i2}^2, \\
 b &= \sum_{i=1}^3 A_i^2 R_{i1} R_{i2}, \\
 c &= \sum_{i=1}^3 A_i^2 R_{i1}^2, \\
 d &= A_1^2 A_2^2 A_3^2 \sum_{i=1}^3 (R_{i3}^2 / A_i^2).
 \end{aligned}$$

Выполняя процедуру, аналогичную описанной выше, т.е. проводя касательную прямую к эллипсу вида (7) в произвольной его точке и принимая во внимание условие перпендикулярности к оси координат x , получим выражение для одномерной проекции эллипсоида на эту ось:

$$h_i = 2\sqrt{c}.$$

Учитывая свойства инвариантности для эллипса вида (7), определим связь между коэффициентами

уравнения (7), осями l_{maxi} и l_{mini} и площадью s_i этого эллипса:

$$\begin{cases}
 s_i = \pi\sqrt{d}; \\
 l_{maxi} = 2\{[(a+c) + [(a+c)^2 - 4d]^{1/2}]/2\}^{1/2}; \\
 l_{mini} = 2\{[(a+c) - [(a+c)^2 - 4d]^{1/2}]/2\}^{1/2}.
 \end{cases} \quad (8)$$

Аналогично можно получить выражения вида (8) для проекции эллипсоида на координатные плоскости xOz и yOz . Исходя из этого определим искомые численные значения базовых признаков проекционных изображений эллипсоида (3) на координатные плоскости:

$$\begin{cases}
 h_i = 2 \left[\sum_{j=1}^3 A_j^2 R_{ji}^2 \right]^{1/2}; \\
 s_i = \pi A_1 A_2 A_3 \left[\sum_{j=1}^3 (R_{ji}^2 / A_j^2) \right]^{1/2}; \\
 l_{maxi} = \\
 = 2 \left\{ \left\{ I_i + \left[I_i^2 - 4A_1^2 A_2^2 A_3^2 \sum_{j=1}^3 (R_{ij}^2 / A_j^2) \right]^{1/2} \right\} / 2 \right\}^{1/2}; \\
 l_{mini} = \\
 = 2 \left\{ \left\{ I_i - \left[I_i^2 - 4A_1^2 A_2^2 A_3^2 \sum_{j=1}^3 (R_{ij}^2 / A_j^2) \right]^{1/2} \right\} / 2 \right\}^{1/2},
 \end{cases} \quad (9)$$

где $I_i = \sum_{j=1}^3 A_j^2 (1 - R_{ji}^2)$; j, i — номера проекции и

текущей ориентировки (ракурса); h_i — длина одномерного проекционного изображения на выбранное направление (ракурс); s_i — площадь двумерного проекционного изображения; l_{maxi} — максимальный линейный размер двумерного проекционного изображения; l_{mini} — минимальный проекционный размер двумерного проекционного изображения.

Таким образом, выражения (9) определяют зависимость линейных (габаритных) размеров трехмерного объекта неправильной формы, аппроксимируемого эллипсоидом общего вида, с численными значениями оптимальных базовых признаков его проекционных изображений и их ракурсами (углами Эйлера).

Полученные базовые признаки используем для определения численных значений геометрических характеристик размеров и формы двумерного проекционного изображения (прагматического изображения) из соотношений

$$D' = 2(s_i/\pi)^{1/2} \text{ или } D' = \frac{l_{maxi} + l_{mini}}{1}; K' = l_{maxi}/l_{mini}$$

где D' и K' — соответственно средний проектируемый диаметр и коэффициент формы двумерного проекционного изображения.

Заключение

Решена прямая задача морфологической обработки изображений трехмерного объекта неправильной формы, аппроксимируемого эллипсоидом общего вида. Полученные при этом математические модели могут быть использованы при разработке, оптимизации и исследовании операций (способов) получения, обработки, анализа и синтеза изображений объектов; при численной оценке качества изображений (по точности, достоверности и быстродействию); при разработке способов реконструкции размеров и формы трехмерных изображений объектов по проекциям; при геометрическом проекционном контроле двумерных и трехмерных объектов; при распознавании образов; в задачах управления и позиционирования робототехники; в задачах верификации программных систем и др.

Список литературы

1. **Беляев С. П., Никифорова Н. К., Смирнов В. В.** Оптико-электронные методы изучения аэрозолей. М.: Энергоиздат, 1983. 385 с.
2. **Доморацкий Е. П.** Метод дистанционного контроля геометрических характеристик движущихся объектов по их дискретным импульсным изображениям // *Материалы II Международной научной конференции "Региональные проблемы дистанционного зондирования Земли"*. Красноярск: Сибирский гос. аэрокосмический университет, 2015. С. 143—146.
3. **Левин Г. Г., Вишняков Г. Н.** Оптическая томография. М.: Радио и связь. 1089. 224 с.
4. **Доморацкий Е. П., Байбикова Т. Н.** О качестве цифровых изображений объектов // *Вестник МФЮА*. М.: МФЮА. 2014. № 2. С. 206—214.
5. **Гонсалес Р., Вудс Р.** Цифровая обработка изображений. М.: Техносфера, 2005. 1042 с.
6. **Методы** компьютерной обработки изображений / Под ред. В. А. Соффера. 2 изд. М.: Физмазлит. 2003. 784 с.
7. **Доморацкий Е. П.** Методика оценки информативности геометрических характеристик проекционных изображений микрообъектов // *Качество. Инновации. Образование*. М.: Известия. 2014. № 2. С. 42—45.
8. **Доморацкий Е. П., Байбикова Т. Н.** О выборе ракурсов проекционных изображений при реконструктивном контроле размеров и формы трехмерных объектов // *Материалы II Международной научной конференции "Региональные проблемы дистанционного зондирования Земли"*. Красноярск: Сибирский гос. аэрокосмический университет, 2015. С. 139—142.
9. **Садовничий Ю. В., Федорчук В. В.** Аналитическая геометрия. Курс лекций с задачами. М.: Экзамен. 2010. 352 с.

E. P. Domorackiy, Professor, e-mail: Domorackiy.E@mfua.ru
Moscow University of Finance and Law (MFUA),
Financial University under the Government of the Russian Federation

Synthesis Algorithm of the Geometrical Characteristics of the Projection Images of Three-Dimensional Objects

The algorithm of synthesis of geometrical characteristics of one-dimensional and two-dimensional projection images of convex three-dimensional objects of irregular shape, which are approximated by a scalene ellipsoid (tri-axial ellipsoid of general form) is under consideration. The algorithm comprises the mathematical models of operation sequences (procedures) for receiving of projection images, for their morphological analysis (receiving their basic geometrical characteristics) and receiving the numerical values of spatial geometrical characteristics of projection images.

Keywords: algorithm, synthesis, geometrical characteristics, object, projection image, mathematical model, approximation, ellipsoid, basic parameters

References

1. **Beljaev S. P., Nikiforova N. K., Smirnov V. V.** *Optiko-elektronnye metody izucheniya ajerorozolej* [Opto-electronic methods for aerosols studying]. Moscow: Jenergoizdat, 1983. 385 p.
2. **Domorackiy E. P.** Metod distancionnogo kontrolja geometricheskix harakteristik dvizhushhihsja ob#ektov po ih diskretnym impul'snym izobrazhenijam [Method of remote control of geometrical characteristics of the moving objects with the usage of their discrete impulse images], *Materialy II Mezhdunarodnoj nauchnoj konferencii "Regional'nye problemy distancionnogo zondirovanija Zemli"*. Krasnojarsk: Sibirskij gos. ajerokosmicheskij universitet, 2015, pp. 143—146.
3. **Levin G. G., Vishnjakov G. N.** *Opticheskaja tomografija* [Optical tomography]. Moscow: Radio i svjaz'. 1089. 224 p.
4. **Domorackiy E. P., Baybikova T. N.** O kachestve cifrovix izobrazhenij ob#ektov [The quality of digital images of objects], *Vestnik MFJuA*. Moscow: MFJuA, 2014, no. 2, pp. 206—214.
5. **Gonsales R., Vuds R.** Cifrovaja obrabotka izobrazhenij [Digital image processing], Moscow: Tehnosfera, 2005. 1042 p.
6. **Metody komp'juternoj obrabotki izobrazhenij** [Methods of computer image processing]. Pod red. V. A. Sojfera. 2 izd. Moscow: Fizmazlit. 2003. 784 p.
7. **Domorackiy E. P.** Metodika ocenki informativnosti geometricheskix harakteristik proekcionnyx izobrazhenij mikroob#ektov [Methodology for assessment of informativeness of the geometrical characteristics of the projection images of micro-objects], *Kachestvo. Innovacii. Obrazovanie*. Moscow: Izdatel'stvo Izvestija, 2014, no. 2, pp. 42—45.
8. **Domorackiy E. P., Baybikova T. N.** O vybore rakurov proekcionnyx izobrazhenij pri rekonstruktivnom kontrole razmerov i formy trehmernyx ob#ektov [About choosing of foreshortenings of the projection images at the reconstructive control of the size and shape of three-dimensional objects], *Materialy II Mezhdunarodnoj nauchnoj konferencii "Regional'nye problemy distancionnogo zondirovanija Zemli"*. Krasnojarsk: Sibirskij gos. ajerokosmicheskij universitet, 2015, pp. 139—142.
9. **Sadovnichij Ju. V., Fedorchuk V. V.** *Analiticheskaja geometrija*. Kurs lekcij s zadachami [Analytical geometry. A course of lectures with tasks]. Moscow: Jekzamen. 2010. 352 p.

УДК 004.272.43

А. Э. Саак, д-р техн. наук, доц., e-mail: saak@tgn.sfedu.ru,
Южный федеральный университет, г. Ростов-на-Дону — Таганрог

О качестве диспетчеризации массивов точных форм с заявками параболического типа в Grid-системах

Проводится анализ функционирования Grid-систем централизованной архитектуры с мультисайтным режимом обслуживания заявки при диспетчеризации массивов точных форм, содержащих заявки параболического типа. Адаптивность рассматриваемых полиномиальных алгоритмов распределения вычислительно-временных ресурсов продемонстрирована на тестовых массивах, индуцированных известными задачами диссекции и облицовки квадрата.

Ключевые слова: Grid-система централизованной архитектуры, мультисайтный режим обслуживания заявки, массив точной формы, заявка кругового типа, заявка гиперболического типа, заявка параболического типа, неэвклидова эвристическая мера, полиномиальная трудоемкость алгоритма

Введение

Grid-системы централизованной архитектуры [1, 2], состоящие из сайтов, содержащих параллельные системы, с мультисайтным диспетчеризованием, характеризующимся возможностью выполнения многопроцессорной заявки (*multiprocessor task, parallel task, parallel job*) на нескольких сайтах одновременно [2, 3], моделируются ресурсным квадрантом [4, 5].

Поступающие заявки пользователей формируются в пакеты (*batch of jobs* [6]), что позволяет онлайн-диспетчеризацию (*on-line scheduling*) свести к ряду офлайн-диспетчеризаций (*off-line scheduling*) [7].

Полагаем, что число процессоров, требуемых в заявке, определяет пользователь, и это число неизменно в процессе выполнения — жесткие заявки (*rigid job*) [8]. Время решения заявки (*processing time*) известно до ее выполнения (*clairvoyant rigid jobs*) [9] и считается целым числом [6, 9—11]. Заявка выполняется без прерывания от начала до окончания (*parallel rigid non-preemptive jobs*) [6, 11]. Для выполнения заявки диспетчер выделяет процессоры с последовательными номерами (*contiguous parallel tasks*) [10, 11]. Сделанные предположения позволяют представлять заявку пользователя при обслуживании диспетчером Grid-системы ресурсным прямоугольником с горизонтальным и вертикальным измерениями, равными соответственно числу единиц ресурса времени и процессоров, требуемому для выполнения заявки [11]. Символом $[a(j), b(j)]$ обозначается j -я заявка, требующая $a(j)$ единиц времени и $b(j)$ единиц процессоров.

Качество диспетчеризации оценивается неэвклидовой эвристической мерой, учитывающей площадь и форму занятой ресурсной области [4]. Эвристическая мера достигает минимума, равного $1/2$, при беспустотной укладке в квадрат.

Постановка задачи

Задача диспетчеризации в Grid-системах централизованной архитектуры с мультисайтным режимом эквивалентна задаче упаковки ресурсных прямоугольников в квадрат с минимальной стороной (*Packing Rectangles into a Square in the Oriented case* (PRSO) [12]). В [12] указывалось на NP-трудность задачи PRSO.

Экспоненциальная сложность оптимального распределения ресурсов требует практического использования эвристических алгоритмов полиномиальной трудоемкости. В качестве аппарата теории диспетчеризации полиномиальной трудоемкости в работах [4, 13—16] предложена среда ресурсных прямоугольников, в которой используются эвристические полиномиальные алгоритмы распределения вычислительно-временных ресурсов, основанные на введенных операциях над ресурсными прямоугольниками.

Для массивов ресурсных прямоугольников, обладающих свойством монотонности горизонтальных или вертикальных измерений, в [16] предложен уровневый алгоритм по высоте с недостатком, в [4] — уровневый алгоритм по протяженности с недостатком, в [17] — уровневый алгоритм по высоте с превышением и уровневый волновой алгоритм (с минимальным отклонением).

Для массивов ресурсных прямоугольников, не обладающих свойством монотонности горизонтальных или вертикальных измерений, в [18] определены V -уровневые алгоритмы по высоте: с недостатком, с превышением, с минимальным отклонением и H -уровневые алгоритмы по протяженности: с недостатком, с превышением, с минимальным отклонением.

Квадратичный тип для массивов из не менее двух заявок определен в [4]. Квадратичный тип одной заявки введен в [19]. Массив ресурсных прямоугольников с беспустотной квадратной ресурсной оболочкой, называемый массивом точной формы, определен в [20]. Качество диспетчирования массивов точной формы из заявок кругового типа исследовалось в работе [20], где тестовым примером служили массивы прямоугольников, индуцированные диссекцией квадрата на различные квадраты. Качество диспетчирования массивов точной формы из заявок кругового и гиперболического типа изучалось в [18], где как тестовый пример использовались массивы прямоугольников, индуцированные диссекцией квадрата на различные прямоугольники с отношением сторон 1:2.

В настоящей статье ставится вопрос об адаптивности полиномиальных алгоритмов для массивов точной формы с заявками параболического типа.

Диспетчеризация уровневymi алгоритмами массива точной формы с заявками параболического типа

Тестовые примеры начнем с диссекции квадрата на последовательные домино и тромино (*sequential domino-tromino dissection*) [21]. Домино называется прямоугольником с отношением большей стороны к меньшей 2:1, тромино — это прямоугольник с отношением большей стороны к меньшей 3:1. Порядком диссекции k назовем число последовательных домино и тромино. В [21] приведены диссекции квадрата 14×14 порядка $k = 6$ и квадрата 17×17 порядка $k = 7$ (рис. 1).

В соответствии с определениями, данными в работе [19], при горизонтальной ориентации квадратичный тип прямоугольника с отношением сторон 2:1 считается круговым, а при вертикальной — гиперболическим. Квадратичный тип прямоугольника с отношением сторон 3:1 считается параболическим [19]. К примеру, на рис. 1 прямоугольник 9×3 является тромино и имеет параболический тип.

Таким образом, массив прямоугольников, индуцированный диссекцией квадрата на последовательные домино и тромино, содержит прямоугольники параболического типа. Массивы ресурсных прямоугольников, индуцированные диссекциями квадрата на последовательные

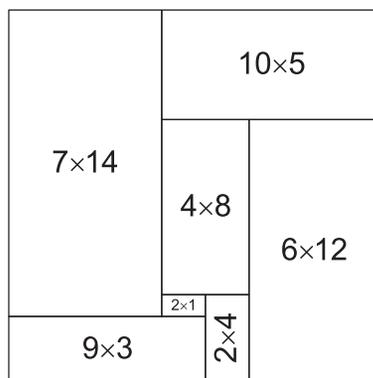


Рис. 1. Диссекция квадрата 17×17 на $k = 7$ последовательных домино, тромино [21]

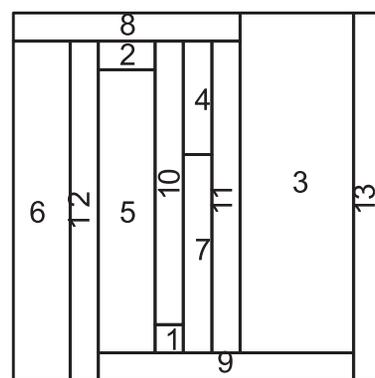


Рис. 2. Диссекция соотношения сторон порядка $k = 13$ [22]

домино и тромино, упорядоченные по убыванию высот, обозначим следующим образом: для $k = 6$ — массив I, для $k = 7$ — массив II.

Следующие тестовые примеры порождены диссекциями соотношения сторон (*aspect ratio dissections*) [22]. Диссекцией соотношения сторон называется диссекция квадрата на прямоугольники с последовательными целыми соотношениями сторон, начиная с единицы [22]. Прямоугольник имеет соотношение сторон j , если отношение большей стороны к меньшей равно j [22]. Порядком диссекции соотношения сторон k назовем максимальное j . В [22] приведены диссекции соотношения сторон порядка $k = 10, 11, 12, 13$ (рис. 2). На прямоугольнике указано значение соотношения сторон.

Как отмечалось выше, прямоугольник с соотношением сторон $j = 2$ относится к круговому или гиперболическому квадратичному типу в зависимости от ориентации. Как отмечалось в [19], квадрат относится к круговому типу. В соответствии с определениями [19], квадратичный тип прямоугольника с соотношением сторон $3, 4, \dots, j, j \geq 3$, считается параболическим. Действительно, условие отнесения прямоугольника к параболическому типу $(b(j) - a(j))^2 > a(j)b(j)$, сформулированное в [19], для прямоугольников с соотношением сторон $j \geq 3$ приводит к очевидному неравенству $(j - 1)^2 > j$ при $j \geq 3$. К примеру, на рис. 2 прямоугольники с соотношением сторон $j = 3, 4, \dots, 13$ имеют параболический тип.

Таким образом, массив прямоугольников, индуцированный диссекцией соотношения сторон, содержит прямоугольники параболического типа. Массивы ресурсных прямоугольников, индуцированные диссекциями соотношения сторон, упорядоченные по убыванию высот, обозначим следующим образом: для $k = 10$ — массив III, $k = 11$ — массив IV, $k = 12$ — массив V, $k = 13$ — массив VI.

Дальнейшие тестовые примеры порождены облицовкой квадрата полосами из меньших квадратов (*tiling squares with strips of smaller squares*) [23]. По-

Эвристические меры ресурсных оболочек V -уровневого алгоритма по высоте с недостатком для массива с параболическими заявками

Номер массива	Эвристическая мера	Номер массива	Эвристическая мера	Номер массива	Эвристическая мера
I	1,03	V	0,84	IX	0,83
II	1,33	VI	1,28	X	0,82
III	0,83	VII	0,91	XI	1,26
IV	1,28	VIII	0,88	XII	1,03

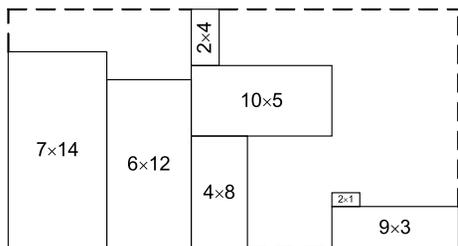


Рис. 5. Укладка массива II V -уровневым алгоритмом по высоте с недостатком

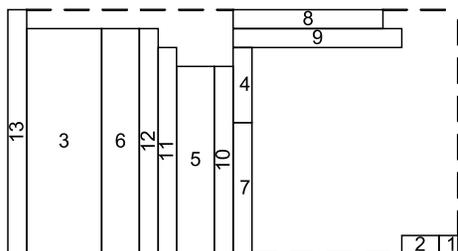


Рис. 6. Укладка массива VI V -уровневым алгоритмом по высоте с недостатком

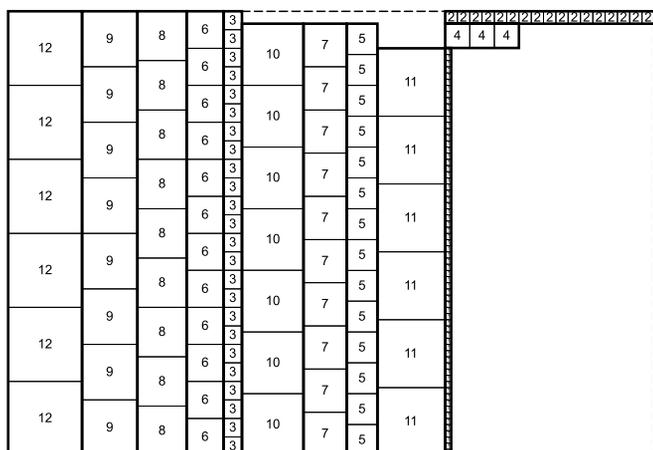


Рис. 7. Укладка массива IX V -уровневым алгоритмом по высоте с недостатком

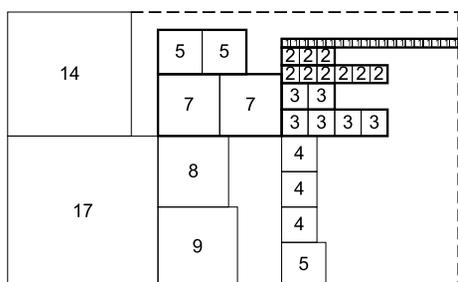


Рис. 8. Укладка массива XII V -уровневым алгоритмом по высоте с недостатком

Видим, что эвристические меры ресурсных оболочек V -уровневого алгоритма по высоте с недостатком для массива с заявками параболического типа не превосходят значения $0,5 + 0,83$.

Построения V -уровневым алгоритмом по высоте с превышением для массивов II, VI, IX, XII приведены на рис. 9–12.

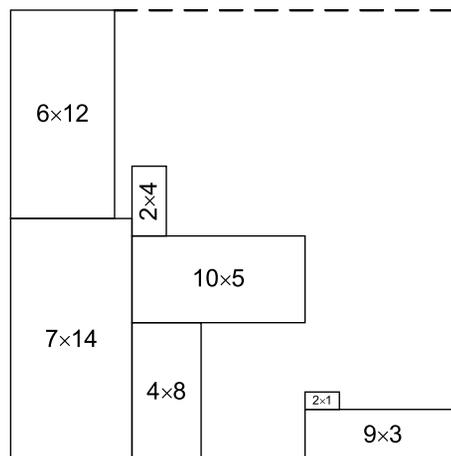


Рис. 9. Укладка массива II V -уровневым алгоритмом по высоте с превышением

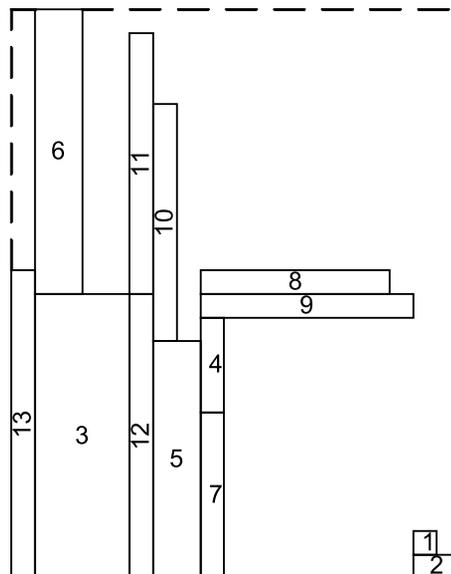


Рис. 10. Укладка массива VI V -уровневым алгоритмом по высоте с превышением

Эвристические меры ресурсных оболочек V -уровневого алгоритма по высоте с превышением для массива с заявками параболического квадратичного типа приведены в табл. 2.

Видим, что эвристические меры ресурсных оболочек V -уровневого алгоритма по высоте с превы-

Таблица 3

Эвристические меры ресурсных оболочек *V*-уровневого алгоритма по высоте с минимальным отклонением для массива с параболическими заявками

Номер массива	Эвристическая мера	Номер массива	Эвристическая мера	Номер массива	Эвристическая мера
I	1,03	V	0,84	IX	0,83
II	0,81	VI	1,28	X	0,82
III	1,64	VII	0,91	XI	1,26
IV	1,28	VIII	0,88	XII	1,03

Эвристические меры ресурсных оболочек *V*-уровневого алгоритма по высоте с минимальным отклонением для массива с заявками параболического квадратичного типа приведены в табл. 3.

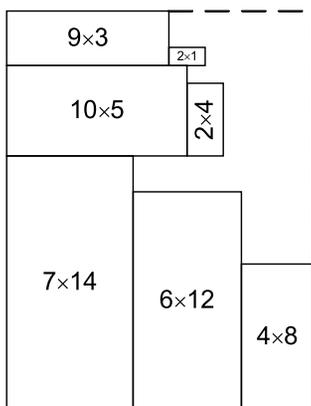


Рис. 17. Укладка массива II *H*-уровневым алгоритмом по протяженности с недостатком

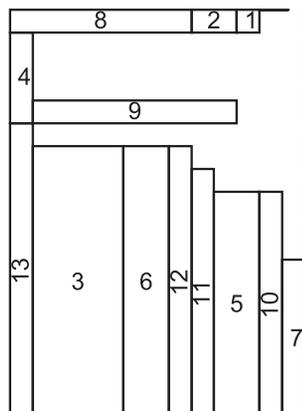


Рис. 18. Укладка массива VI *H*-уровневым алгоритмом по протяженности с недостатком

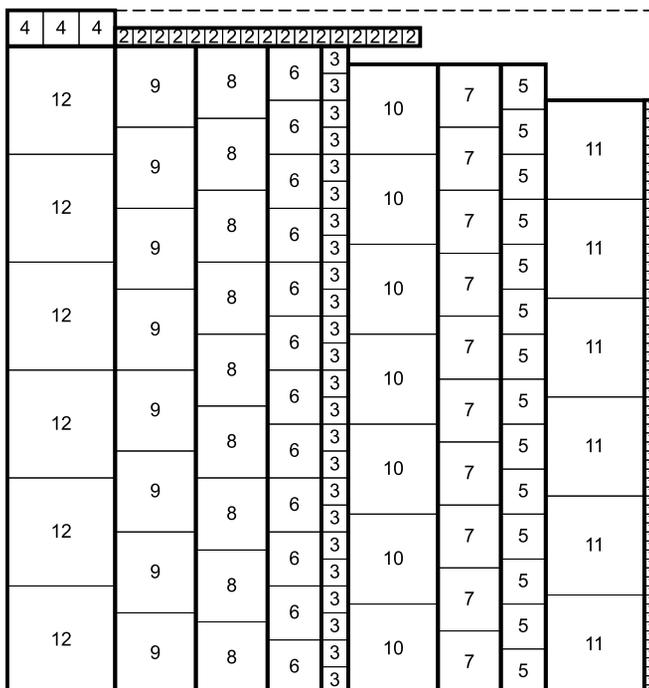


Рис. 19. Укладка массива IX *H*-уровневым алгоритмом по протяженности с недостатком

Таблица 4

Эвристические меры ресурсных оболочек *H*-уровневого алгоритма по протяженности с недостатком для массива с параболическими заявками

Номер массива	Эвристическая мера	Номер массива	Эвристическая мера	Номер массива	Эвристическая мера
I	1,01	V	0,84	IX	0,53
II	0,69	VI	0,77	X	0,64
III	1,24	VII	0,56	XI	0,56
IV	0,91	VIII	0,54	XII	0,59

Видим, что эвристические меры ресурсных оболочек *V*-уровневого алгоритма по высоте с минимальным отклонением для массива с заявками параболического типа не превосходят значения $0,5 + 1,14$.

Построения *H*-уровневым алгоритмом по протяженности с недостатком для массивов II, VI, IX, XII приведены на рис. 17—20.

Эвристические меры ресурсных оболочек *H*-уровневого алгоритма по протяженности с недостатком для массива с заявками параболического квадратичного типа приведены в табл. 4.

Видим, что эвристические меры ресурсных оболочек *H*-уровневого алгоритма по протяженности с недостатком для массива с заявками параболического типа не превосходят значения $0,5 + 0,74$.

Построения *H*-уровневым алгоритмом по протяженности с превышением для массивов II, VI, IX, XII приведены на рис. 21—24.

Эвристические меры ресурсных оболочек *H*-уровневого алгоритма по протяженности с превышением для массива с заявками параболического квадратичного типа приведены в табл. 5.

Видим, что эвристические меры ресурсных оболочек *H*-уровневого алгоритма по протяженности с превышением для массива с заявками параболического типа не превосходят значения $0,5 + 0,51$.

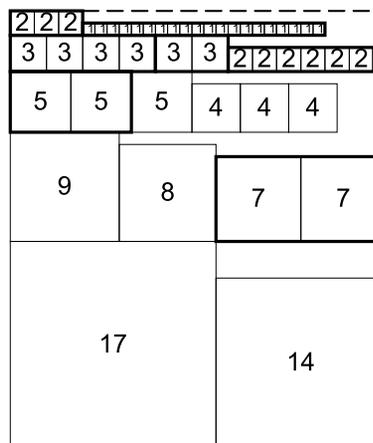


Рис. 20. Укладка массива XII *H*-уровневым алгоритмом по протяженности с недостатком

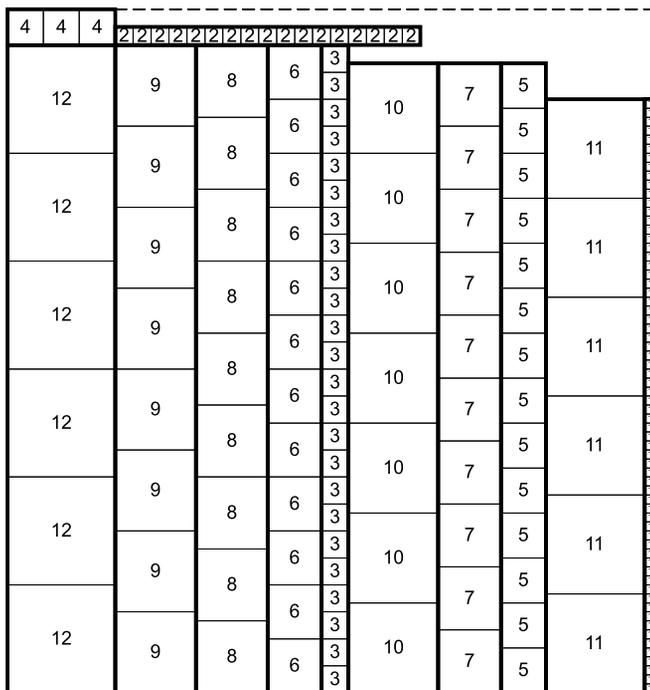


Рис. 27. Укладка массива IX *H*-уровневым алгоритмом по протяженности с минимальным отклонением

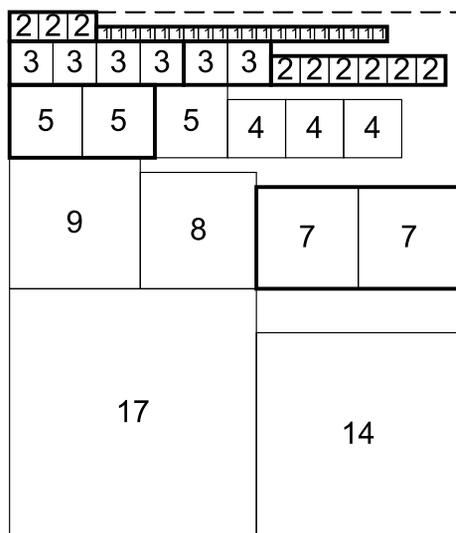


Рис. 28. Укладка массива XII *H*-уровневым алгоритмом по протяженности с минимальным отклонением

Таблица 6

Эвристические меры ресурсных оболочек *H*-уровневого алгоритма по протяженности с минимальным отклонением для массива с параболическими заявками

Номер массива	Эвристическая мера	Номер массива	Эвристическая мера	Номер массива	Эвристическая мера
I	0,7	V	0,81	IX	0,53
II	0,73	VI	0,77	X	0,64
III	0,86	VII	0,59	XI	0,56
IV	0,93	VIII	0,54	XII	0,59

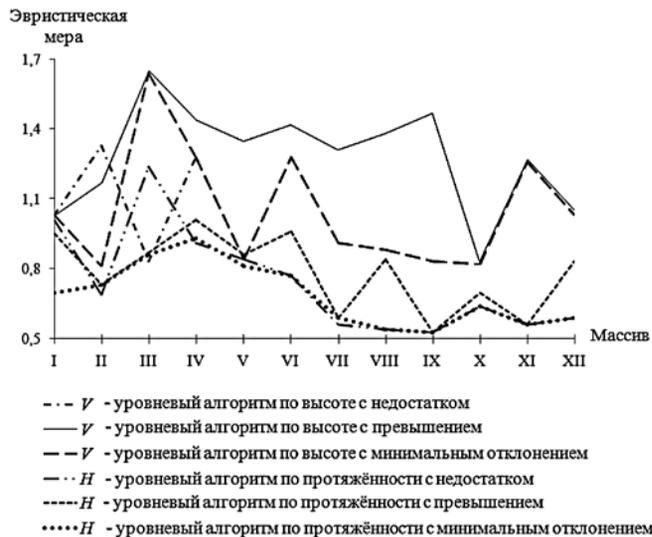


Рис. 29. Эвристические меры ресурсных оболочек *V*-уровневых и *H*-уровневых алгоритмов при диспетчеризации массивами с параболическими заявками

Графики эвристической меры ресурсных оболочек *V*-уровневыми алгоритмами по высоте (с недостатком, с превышением, с минимальным отклонением) и *H*-уровневыми алгоритмами по протяженности (с недостатком, с превышением, с минимальным отклонением) при диспетчеризации массивами с параболическими заявками I–XII показаны на рис. 29.

Видим, что *H*-уровневые алгоритмы по протяженности имеют меньшую эвристическую меру ресурсных оболочек по сравнению с *V*-уровневыми алгоритмами по высоте. Так, наименьшее значение эвристической меры *H*-уровневых алгоритмов по протяженности равно 0,53, тогда как у *V*-уровневых алгоритмов по высоте наименьшее значение эвристической меры равно 0,81. При этом *H*-уровневый алгоритм по протяженности с минимальным отклонением имеет наименьшее значение максимума эвристической меры $0,5 + 0,43$ на рассматриваемых тестовых массивах ресурсных прямоугольников. Исследуемые полиномиальные алгоритмы могут быть использованы в Grid-системах с централизованной структурой при диспетчеризации массивов точных форм с заявками параболического типа.

Заключение

Для массивов точных форм с заявками параболического типа вычисляются эвристические меры ресурсных оболочек различных полиномиальных уровней алгоритмов. Адаптированность анализируемых алгоритмов продемонстрирована на 12 тестовых массивах, индуцированных известными задачами диссекции и облицовки квадрата. При обслуживании массивов точных форм с заявками параболического типа в Grid-системах рекомендуется применять рассматриваемые полиномиальные алгоритмы диспетчеризования.

Список литературы

1. **Rahman M., Ranjan R., Buuya R., Benatallah B.** A taxonomy and survey on autonomic management of applications in grid computing environments // *Concurrency Computat.: Pract. Exper.* 2011. Vol. 23, Iss. 16. P. 1990–2019.
2. **Hamscher V., Schwiegelshohn U., Streit A., Yahyapour R.** Evaluation of job-scheduling strategies for grid computing // *LNCS*. 2000. Vol. 1971. P. 191–202.
3. **Soumez O., Mohamed H., Epema D.** On the benefit of processor coallocation in multicluster grid systems // *IEEE Trans. on Paral. and Distrib. Syst.* 2010. Vol. 21, N. 6, P. 778–789.
4. **Саак А. Э.** Полиномиальные алгоритмы распределения ресурсов в Grid-системах на основе квадратичной типизации массивов заявок // *Информационные технологии*. 2013. № 7. Приложение. 32 с.
5. **Саак А. Э.** Управление ресурсами и заявками пользователей в Grid-системах с централизованной архитектурой // *Труды XII Всероссийского совещания по проблемам управления ВСПУ—2014*. Москва, 16–19 июня 2014 г. М.: Институт проблем управления им. В. А. Трапезникова РАН, 2014. С. 7489–7498.
6. **Bougeret M., Dutot P., Trystram D., Jansen K., Robenek C.** Improved approximation algorithms for scheduling parallel jobs on identical clusters // *Theoretical Computer Science*. 2015. Vol. 600. P. 70–85.
7. **Switalski P., Sredynski F.** Scheduling parallel batch jobs in grids with evolutionary metaheuristics // *J. Sched.* 2015. Vol. 18, Iss. 4. P. 345–357.
8. **Feitelson D., Rudolph L., Schwiegelshohn U., Sevcik K., Wong P.** Theory and practice in parallel job scheduling // *LNCS*. 1997. Vol. 1291. P. 1–34.
9. **Bougeret M., Dutot P., Jansen K., Otte C., Trystram D.** A Fast 5/2-approximation algorithm for hierarchical scheduling // *LNCS*. 2010. Vol. 6271. P. 157–167.
10. **Bładek I., Drozdowski M., Guinand F., Schepler X.** On contiguous and non-contiguous parallel task scheduling // *J. Sched.* 2015. Vol. 18, Iss. 5. P. 487–495.
11. **Caramia M., Giordani S., Iovanello A.** Grid scheduling by on-line rectangle packing // *Networks*. 2004. Vol. 44, N. 2. P. 106–119.
12. **Martello S., Monaci M.** Models and algorithms for packing rectangles into the smallest square // *Computers & Operations Research*. 2015. Vol. 63. P. 161–171.
13. **Саак А. Э.** Локально-оптимальные ресурсные распределения // *Информационные технологии*. 2011. № 2. С. 28–34.
14. **Саак А. Э.** Алгоритмы диспетчеризации в Grid-системах на основе квадратичной типизации массивов заявок // *Информационные технологии*. 2011. № 11. С. 9–13.
15. **Саак А. Э.** Диспетчеризация в Grid-системах на основе однородной квадратичной типизации массивов заявок пользователей // *Информационные технологии*. 2012. № 4. С. 32–36.
16. **Саак А. Э.** Сравнительный анализ полиномиальных алгоритмов диспетчеризации в Grid-системах // *Информационные технологии*. 2012. № 9. С. 28–32.
17. **Саак А. Э.** Уровневые алгоритмы диспетчеризации массивами заявок кругового типа в Grid-системах // *Известия ЮФУ. Технические науки*. 2015. № 6 (167). С. 223–231.
18. **Саак А. Э.** Диспетчеризация массивов заявок кругового и гиперболического типа в Grid-системах // *Информационные технологии*. 2016. № 5. С. 323–332.
19. **Саак А. Э.** Диспетчеризация заявок кругового типа в Grid-системах // *Информационные технологии*. 2016. № 1. С. 37–41.
20. **Саак А. Э., Курейчик В. В.** О качестве диспетчеризации линейных полиэдров точных форм // *Известия ЮФУ. Технические науки*. 2015. № 4 (165). С. 56–67.
21. **Pegg, ed.** Domino-tromino problem. 2001. URL: <http://www.mathpuzzle.com/4Dec2001.htm>
22. **Friedman E.** Problem of the Month (December 2002). Aspect ratio dissections. 2002. URL: <http://www2.stetson.edu/~efriedma/mathmagic/1202.html>
23. **Friedman E.** Problem of the Month (December 2001). Tiling squares with strips of smaller squares. 2015. URL: <http://www2.stetson.edu/~efriedma/mathmagic/1201.html>
24. **Friedman E.** Problem of the Month (November 2003). Hanoi square tiling. 2005. URL: <http://www2.stetson.edu/~efriedma/mathmagic/1103.html>

A. E. Saak, PhD, Head of Department, e-mail: saak@tgn.sfedu.ru
Southern Federal University, Taganrog

To Quality of Scheduling of Precise Form Sets of Parabolic Type Tasks in Grid Systems

Grid systems with centralized architecture, which consist of sites containing parallel systems with multisite scheduling that is characterized by the possibility of multiprocessor task performance on several sites simultaneously, are modeled by resource quadrants. Incoming user's tasks are formed in batches of jobs that allows to turn on-line scheduling into a set of off-line scheduling cases. We suppose that the number of required processors for a task is defined by a user at the moment of feeding it in a system and this number is invariable during the process of task fulfillment — rigid jobs. Processing time is known feature at the beginning, and it is integer. A task could be described as: clairvoyant rigid job; parallel rigid non-preemptive job; contiguous parallel task. The assumptions allow representing user's task during its processing by Grid system's scheduler in the form of resource rectangle with the horizontal and vertical dimensions equaled, correspondingly, to the number of resource units of time and processors required for task performance. Scheduling quality is estimated by the Non-Euclidean heuristic measure which takes into consideration both the area and the form of occupied resource region. The heuristic measure reaches its minimum at in-square packing without emptiness. Thus, the scheduling problem in the Grid systems with centralized architecture and multisite mode is equivalent to the problem of Packing Rectangles into a Square in the Oriented case (PRSO). Exponential complexity of optimal resource distribution demands practical use of heuristic algorithms of polynomial completeness. The author suggests resource rectangles environment as the instrumentation of polynomial completeness scheduling theory. In resource rectangles environment it is suggested the heuristic polynomial algorithms for processor and time resource distribution, which stay on the base of introduced operations on resource rectangles. The problem posed in the paper concerns polynomial algorithms adaptiveness for precise form sets of parabolic type tasks.

Keywords: Grid systems with centralized architecture, multisite mode of task processing, precise form set, task of circular type, task of hyperbolic type, task of parabolic type, Non-Euclidean heuristic measure, Packing Rectangles into a Square in the Oriented case, polynomial completeness of an algorithm, polynomial algorithms for processor and time resource distribution

References

1. **Rahman M., Ranjan R., Buyya R., Benattallah B.** A taxonomy and survey on autonomic management of applications in grid computing environments, *Concurrency Computat.: Pract. Exper.*, 2011, vol. 23, Iss. 16, pp. 1990–2019.
2. **Hamscher V., Schwiigelshohn U., Streit A., Yahyapour R.** Evaluation of job-scheduling strategies for grid computing, *LNCS*, 2000, vol. 1971, pp. 191–202.
3. **Sonmez O., Mohamed H., Epema D.** On the benefit of processor coallocation in multicluster grid systems, *IEEE Trans. on Paral. and Distrib. Syst.*, 2010, vol. 21, no. 6, pp. 778–789.
4. **Saak A. Eh.** Polinomialnye algoritmy raspredeleniya resursov v Grid-sistemakh na osnove kvadrachnoy tipizatsii massivov zayavok [Polynomial algorithms for resource allocation in Grid-based systems for quadratic typing, arrays applications], *Informacionnye tekhnologii* [Information Technologies], 2013, no. 7, Prilozhenie, 32 p.
5. **Saak A. Eh.** Upravleniye resursami i zayavkami polzovatelye v Grid-sistemakh s centralizovannoy arkhitektyroy [Resource and multiprocessor task management in Grid system of centralized architecture], *Trudy XII Vserossiyskogo soveshanya po problem upravleniya VSPU—2014. Moskva, 16 ijunya—19 ijunya 2014 g.* [Proceedings of XII all-Russian conference "Control problems" RCCP'2014. Moscow, 16 June — 19 June 2014]. Moscow: Institut problem upravleniya im. V. A. Trapeznikova RAN, 2014, pp. 7489–7498.
6. **Bougeret M., Dutot P., Trystram D., Jansen K., Robenek C.** Improved approximation algorithms for scheduling parallel jobs on identical clusters, *Theoretical Computer Science*, 2015, vol. 600, pp. 70–85.
7. **Switalski P., Sredynski F.** Scheduling parallel batch jobs in grids with evolutionary metaheuristics, *J. Sched.*, 2015, vol. 18, Iss. 4, pp. 345–357.
8. **Feitelson D., Rudolph L., Schwiigelshohn U., Sevcik K., Wong P.** Theory and practice in parallel job scheduling, *LNCS*, 1997, vol. 1291, pp. 1–34.
9. **Bougeret M., Dutot P., Jansen K., Otte C., Trystram D.** A Fast 5/2-approximation algorithm for hierarchical scheduling, *LNCS*, 2010, vol. 6271, pp. 157–167.
10. **Blądek I., Drozdowski M., Guinand F., Schepler X.** On contiguous and non-contiguous parallel task scheduling, *J. Sched.*, 2015, vol. 18, Iss. 5, pp. 487–495.
11. **Caramia M., Giordani S., Iovanella A.** Grid scheduling by on-line rectangle packing, *Networks*, 2004, vol. 44, no. 2, pp. 106–119.
12. **Martello S., Monaci M.** Models and algorithms for packing rectangles into the smallest square, *Computers & Operations Research*, 2015, vol. 63, pp. 161–171.
13. **Saak A. Eh.** Lokalno-optimalnye resursnye raspredeleniya [Locally optimal resource allocation], *Informatsionnye tekhnologii* [Information Technologies], 2011, no. 2, pp. 28–34.
14. **Saak A. Eh.** Algoritmy dispetcherizatsii v Grid-sistemakh na osnove kvadrachnoy tipizatsii massivov zayavok [Algorithms scheduling in Grid-based systems for quadratic typing, arrays applications], *Informatsionnye tekhnologii* [Information Technologies], 2011, no. 11, pp. 9–13.
15. **Saak A. Eh.** Dispetcherizatsiya v GRID-sistemakh na osnove odnorodnoy kvadrachnoy tipizatsii massivov zayavok polzovatelye [Scheduling in GRID-systems on the basis of homogeneous quadratic typing, arrays of user requests], *Informatsionnye tekhnologii* [Information Technologies], 2012, no. 4, pp. 32–36.
16. **Saak A. Eh.** Sravnitelniy analiz polinomialnykh algoritmov dispetcherizatsii v GRID-sistemakh [Comparative analysis of polynomial algorithms for scheduling in GRID-systems], *Informatsionnye tekhnologii* [Information Technologies], 2012, no. 9, pp. 28–32.
17. **Saak A. Eh.** Urovnevye algoritmy dispetcherizatsii massivamy zayavok krygovogo tipa v Grid-sistemakh [Level algorithms of scheduling by circle type task sets in grid systems], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, no. 6 (167), pp. 223–231.
18. **Saak A. Eh.** Dispetcherizatsiya massivov zayavok krygovogo i giperbolicheskogo tipa v Grid-sistemakh [Scheduling of sets of circular-type and hyperbolic-type tasks in grid-systems], *Informatsionnye tekhnologii* [Information Technologies], 2016, vol. 22, no. 5, pp. 323–332.
19. **Saak A. Eh.** Dispetcherizatsiya zayavok krygovogo tipa v Grid-sistemakh [Circular-typed multiprocessor tasks scheduling in grid systems], *Informatsionnye tekhnologii* [Information Technologies], 2016, vol. 22, no. 1, pp. 37–41.
20. **Saak A. Eh., Kureichik V. V.** O kachestve dispetcherizatsii lineynix poliedralnykh tochnix form [To quality of precisely formed linear polyedrals scheduling], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2015, no. 4 (165), pp. 56–67.
21. **Pegg, ed.** Domino-tromino problem. 2001. URL: <http://www.mathpuzzle.com/4Dec2001.htm>
22. **Friedman E.** Problem of the Month (December 2002). Aspect ratio dissections. 2002. URL: <http://www2.stetson.edu/~efriedma/mathmagic/1202.html>
23. **Friedman E.** Problem of the Month (December 2001). Tiling squares with strips of smaller squares. <http://www2.stetson.edu/~efriedma/mathmagic/1201.html>
24. **Friedman E.** Problem of the Month (November 2003). Hanoi square tiling. 2015. URL: <http://www2.stetson.edu/~efriedma/mathmagic/1103.html>

БЕЗОПАСНОСТЬ ИНФОРМАЦИИ CRYPTOSAFETY INFORMATION

УДК 004.056.5; 681.142.342

Э. М. Димов, д-р техн. наук, проф.,

О. Н. Маслов, д-р техн. наук, проф., зав. каф., e-mail: maslov@psati.ru,

А. С. Раков, канд. техн. наук, докторант,

Поволжский государственный университет телекоммуникаций и информатики, г. Самара

Управление информационной безопасностью корпорации с применением критериев риска и ожидаемой полезности

Рассматривается проблема конвергенции принципов и количественных критериев теорий знаний, риска и ожидаемой полезности. Конвергенция позволяет использовать критерии приемлемого риска и ожидаемой полезности при статистическом имитационном моделировании (СИМ) по методу Димова — Маслова (МДМ). Целью СИМ по МДМ является исследование и управление сложных систем (СС) организационно-технического типа. Примером СС организационно-технического типа является система обеспечения информационной безопасности (СОИБ) корпорации. С помощью компьютерной технологии метода Монте-Карло решается главная проблема моделирования — уменьшение влияния неопределенности исходных и промежуточных данных на результаты СИМ по МДМ. Показана эффективность уменьшения неопределенности выбора управленческих решений при проектировании СОИБ. Исходные данные для проведения СИМ по МДМ формируются на основе верифицированных и аксиологических знаний об объекте. Для выбора управленческих решений сформулирован критерий оправданного риска, который соответствует функционалу ожидаемой полезности. Представлена схема формирования функционала ожидаемой полезности при проведении СИМ по МДМ. Рассмотрены особенности бизнес-процесса разработки СОИБ корпорации в виде системы активной защиты (САЗ) конфиденциальной информации (КИ) коммерческого назначения. Представлена концептуальная логическая схема бизнес-процесса разработки САЗ КИ. Намечены пути анализа эффективности применения СОИБ корпорации с помощью СИМ по МДМ.

Ключевые слова: управление корпорацией, информационная безопасность, статистическое имитационное моделирование, риск, ожидаемая полезность

Введение

"Подлинный прогресс в любой науке наступал тогда, когда в ходе изучения задач, которые были скромными по сравнению с окончательными целями, развивались методы, которые можно было обобщать все дальше и дальше" [1]. Правомерность этого утверждения иллюстрируют современные теории риска и ожидаемой полезности, которые поначалу решали частные задачи в области азартных игр и лотерей, чтобы затем перейти к анализу экономических, экологических, военных и других проблем, актуальных и значимых поистине в глобальном масштабе. С точки зрения теории сложных систем (СС) важно также, что эти теории являются открытыми СС в том смысле, что допускают конвергенцию с методами и средствами, принадлежащими другим областям знаний — в частности, связанными с новыми информационными технологиями.

Под конвергенцией (от лат. *convergere* — приближаться, сходиться) будем понимать развитие СС по пути сближения и приобретения общих или

сходных признаков. Примером новых информационных технологий является метод статистического имитационного моделирования (СИМ) по версии Димова—Маслова (МДМ) [2—3], включающий методы сценариев и функционально-стоимостного анализа, а также использующий технологию метода Монте-Карло (ММК). Стимулом для указанной конвергенции являются синергетические эффекты, многократно увеличивающие ценность знаний, присущих данным теориям.

Практический интерес для СИМ по МДМ сегодня в большей степени представляют онтологические принципы развития теорий риска и ожидаемой полезности, нежели их аналитический аппарат и математические модели [4]. Ключевым фактором при этом становится неопределенность знаний [5], сопровождающая процесс моделирования на всех этапах его реализации: от определения исходных данных, параметров и переменных до выбора критериев оценки полученных результатов. Механизм уменьшения неопределенностей приобретает важнейшее значение — во-первых, поскольку приме-

нение точных формул к неточным величинам есть "обман и пустая трата времени" [6]. Во-вторых, потому что мера, а также само понятие неопределенности — применительно к решению конкретных задач — каждый раз нуждаются в уточнении и доопределении.

Цель статьи — анализ перспектив и возможностей, которые дает применение критериев риска и ожидаемой полезности при исследовании и управлении СС с использованием СИМ по МДМ, на примере проектирования фрагмента системы обеспечения информационной безопасности (СОИБ) корпорации [7, 8], предназначенного для активной защиты конфиденциальной информации от утечки во внешнюю среду через случайные антенны [9—10].

Онтологическая модель ситуации при проведении СИМ

При управлении СС организационно-технического типа (экономическими, экологическими, информационными, военными [11] и т. п.) находят применение экспертные (эвристические), вероятностно-теоретические, вероятностно-статистические и статистические модели и методы. В совокупности они образуют концептуальную платформу (или, по другому названию, — онтологическую модель ситуации (ОМС) [12]) как основу для принятия управленческих решений. В состав ОМС входят персональные (индивидуальные) онтологии (от греч. *ontos* — сущее), основанные на собственных знаниях лиц, принимающих решения (ЛПР), и групповые онтологии, которые формируются путем договоренности между ними, включающие как верифицированные знания, признаваемые всеми ЛПР, так и аксиологические знания (предзнания в виде убеждений, жизненного опыта и гипотез), остающиеся предметом дискуссий. Присутствие ЛПР является неотъемлемым признаком организационно-технических СС, поскольку их свойства и поведение во многом объясняются наличием этого "слабоструктурируемого и плохоформализуемого" человеческого фактора [13].

Характерным примером организационно-технической СС является СОИБ, в которую входят техническая часть (в виде оборудования для защиты конфиденциальной информации), а также организационная составляющая, обусловленная в том числе взглядами ЛПР на процесс функционирования СОИБ, способы контроля и защиты конфиденциальной информации. Применительно к проектированию и исследованию СОИБ экспертные методы представляются слишком "слабыми", а статистические — слишком "жесткими". При этом метод СИМ по МДМ сочетает достоинства теоретико-вероятностного и статистически-вероятностного подходов, а также компьютерного варианта технологии ММК [2]. Его отличительными особенностями являются:

- системный подход — ввиду многообразия свойств и рабочих характеристик СС, а также нестандартных и не всегда однозначных требований к их статическим и динамическим моделям;
- ориентация на управление бизнес-процессами в корпоративных СС, в том числе связанными с управлением СОИБ;
- использование знаний разного типа (верифицированных и аксиологических) при определении исходных данных и других условий для проведения СИМ;
- доступность практической реализации.

Теории риска и ожидаемой полезности для оценки неопределенности используют линейный критерий вида

$$R_S = P(A)A_S,$$

где $P(A)$ — вероятность результата A функционирования СС; A_S — его полезность (стоимость связанного с риском выигрыша или проигрыша в выбранных единицах); R_S для краткости также именуется просто риском. Данный критерий является базовым при проектировании СОИБ и тесно связан с понятием неопределенности знаний. При этом, согласно [8—10], следует различать неопределенности, которые относятся к реальным СС, и неопределенности, сопровождающие проектирование СС, поскольку в первом случае речь идет о действиях ЛПР в материальной (реальной) среде, а во втором случае — в виртуальной среде моделирования СС. Соответственно, в первом случае работает концепция "риск как неопределенность", где значение R_S считается объективно существующей случайной величиной, связанной с вероятностным распределением возможных исходов некоторой операции (процедуры, процесса) на объекте СИМ. Во втором случае в рамках концепции "риск как возможность" значение R_S может соответствовать однократному негативному событию A , которого хотелось бы избежать совсем, а значение $P(A)$ представляет собой часть вероятности, характеризующей полную группу событий, включающей A , которая, по мнению эксперта или группы экспертов, принимающих решение, относится к событию A . Данный подход полностью соответствует идеологии СИМ по МДМ и с успехом может быть использован при проектировании СОИБ.

Выделим три варианта моделирования неопределенности X в зависимости от объема имеющейся у ЛПР информации о параметрах и свойствах объекта СИМ:

- моделирование при максимуме неопределенности, когда можно воспользоваться "принципом безразличия" и полагать, что все возможности или варианты развития событий являются равновероятными, что соответствует равномерно распределению X ;

- моделирование ситуации, соответствующей условиям применимости предельных теорем теории вероятностей, когда моделью неопределенности можно считать плотность одномерного устойчивого закона, частным случаем которого является нормальный закон [9];
- моделирование при минимуме неопределенности, который соответствует "квазидетерминированной" модели с малой дисперсией вплоть до дельта-функции Дирака.

Пусть ЛПР S_n , принадлежащее совокупности n [1; M], рассматривает K вариантов своих действий (сценариев развития событий), направленных на достижение поставленной цели как некоего важного для него позитивного эффекта (выигрыша) F_k , который может быть им достигнут с вероятностью p_F при затратах G_k , которые могут быть обеспечены им с вероятностью p_G . При этом S_n полагает, что k -й сценарий приводит к достижению цели с вероятностью p_k , которая характеризует некие объективные (зависящие не только от него лично) внешние и внутренние обстоятельства. Тогда в качестве критерия эффективности k -го сценария S_n может быть выбрана расчетная величина

$$Q_k = p_k(p_F F_k - p_G G_k), \quad (1)$$

где k [1; K], которая и является при проектировании СОИБ оправданной (по мнению ЛПР) мерой риска — по смыслу данный критерий близок к объективной ожидаемой полезности [1]. Но объективно существующими величинами в (1) можно считать p_k ; p_G и G_k , тогда как вероятность p_F получения выигрыша и его величина F_k являются виртуальными и, стало быть, могут оцениваться лишь субъективными (эвристическими, экспертными) методами. В связи с этим (1) одновременно является критерием субъективной ожидаемой полезности [14], что обусловлено природой рассматриваемой СС организационно-технического типа. Кроме того, вероятности p_G , p_k и p_F , а также выигрыш F_k в (1) достаточно сложным образом зависят от G_k , поскольку пока субъект S_n не пойдет на затраты и не обеспечит ресурсы, необходимые хотя бы для начала проекта, получение выигрыша для него маловероятно, тогда как по мере увеличения $p_G G_k$ эта вероятность растет и т. д.

Подчеркивая вероятностную сущность критерия Q_k , условимся именовать его функционалом ожидаемой полезности (ФОП) [15]. Отметим, что ФОП соответствует наиболее очевидной и простой модели неопределенности знаний ЛПР, но в то же время, во-первых, он отражает стохастическую сущность риска, который вместо категорий "да" и "нет" оперирует понятиями "может быть", "скорее всего", что устраняет недостатки и противоречия, присущие детерминистским подходам. Во-вторых, ФОП полностью соответствует идеологии и методике СИМ по МДМ, что особенно важно для при-

ложений, где исследование СС другими способами (аналогично разработке и оценке эффективности будущего применения СОИБ) практически невозможно.

Моделирование функционала ожидаемой полезности

Двухэтапная схема моделирования процесса формирования ФОП Q_k для проекта, которая соответствует четырем возможным сценариям его реализации, представлена на рис. 1. На первом этапе, когда процесс является преимущественно расходящимся, он может быть описан с помощью логических связей типа дизъюнкции "ИЛИ", поскольку на каждом шаге процесса точка разветвления, принадлежащая верхнему уровню, порождает не менее двух таких же точек на соседнем нижнем уровне (см. уровни 1 и 2 на схеме рис. 1).

Аналогичным образом на втором этапе, где процесс является преимущественно сходящимся, для его описания больше подходят связи типа конъюнкции "И" — объединяющие точки разветвления, как это демонстрируют уровни K ; L и N на схеме рис. 1. Трехточечные фрагменты рассматриваемого процесса, выделенные на рис. 1 штриховыми контурами, позволяют исследовать типовые элементы процесса формирования Q_k , для моделирования которых, наряду с методом сценариев, будем использовать метод функционально-стоимостного анализа [8], а также другие методы и модели, используемые

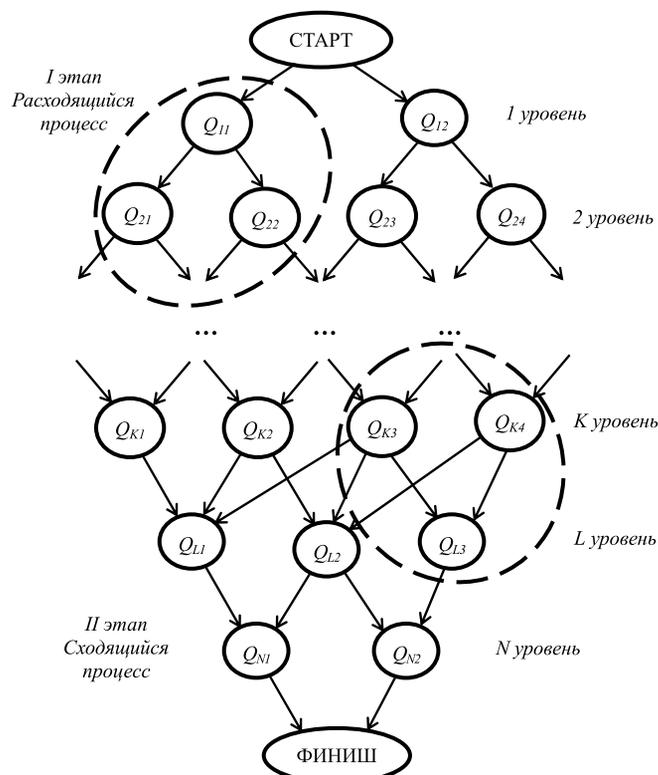


Рис. 1. Двухэтапная схема процесса формирования ФОП

при исследовании СС [2—3], так как возможностей теории вероятностей для описания Q_k на всех шагах от старта (начала) до финиша (завершения) процесса на рис. 1 может быть недостаточно.

Согласно рис. 1, на старте процесса ЛПР, даже "проигрывая" мысленно весь ход развития дальнейших событий в каждом из четырех вариантов, должен сделать выбор между двумя конкретными шагами (действиями), которые связаны для него с промежуточными полезностями Q_{11} и Q_{12} , для чего ему необходимо оценить формирующие их величины p_F , p_G , F_k и G_k . Оценка выигрыша F_k и затрат G_k с помощью функционально-стоимостного анализа может быть как точной (например, в денежном выражении), так и условной (в процентах, баллах или других единицах — см. пример далее), способ этой оценки на разных шагах процесса может быть разным, однако важно, во-первых, чтобы он не менялся на уровне данного рассматриваемого шага. Во-вторых, при невозможности вычислить F_k и G_k в детерминированном виде они могут быть найдены как случайные числовые величины, распределенные в заданных пределах по известным или неизвестным ЛПР вероятностным законам, а также "гибридными" способами (интервальный анализ, метод нечетких множеств и т. п.). Если от значений F_k и G_k на данном шаге далее ничего не зависит, их можно принять равными любой отличной от нуля константе (например, единице). Наличие в схеме на рис. 1 одной финишной точки соответствует процессу с единственным выходом. Примером могут быть сценарии действий ЛПР (хакера, потенциального злоумышленника, недобросовестного конкурента), направленные на достижение поставленной негативной цели [8], учитываемые при проектировании СОИБ.

При оценке фигурирующих в составе Q_{11} и Q_{12} вероятностей p_F и p_G экспертными методами их субъективные значения также можно считать случайными числовыми величинами или находить "гибридными" способами, аналогично F_k и G_k , но следуя формальным правилам теории вероятностей для обеспечения идентичности и симметрии проводимого анализа в каждом из рассматриваемых вариантов — во избежание ошибки на финише, при выборе ЛПР между итоговыми полезностями Q_{N1} и Q_{N2} .

Следует также учитывать, что процесс на рис. 1 представляет собой СС нерелефторного типа, поэтому на него не распространяется принцип максимума Л. С. Понтрягина [2]. Это означает, что путем выбора максимальных значений промежуточной полезности Q_k на каждом уровне формирования заключительной полезности невозможно "смонтировать" оптимальный процесс, добившись условия $Q_N = \text{MAX}(\text{MAX})$. Особенно наглядно это проявляется в случае, когда цепочки шагов на рис. 1 реализуются по каузальной схеме и "перескочить" с одной ветви на другую по ходу процесса нельзя.

Поэтому каждый вариант следует исследовать от начала до конца и давать ему оценку в законченном виде, что не исключает возможности разветвлять и объединять ветви, а также формировать дополнительные варианты реализации процесса с учетом полученной по результатам моделирования новой информации о нем.

При рассмотрении схемы на рис. 1 с точки зрения теории управления СС возникают два вопроса: во-первых, существует ли квазиоптимальный (лучший из числа доступных) по критерию Q_N вариант реализации процесса, и, во-вторых, располагают ли ЛПР корпоративными ресурсами (временными, материально-техническими, финансовыми, интеллектуальными) для воплощения данного варианта в жизнь. В обоих случаях речь идет о максимально быстром и безошибочном "просчете" характеристик шагов во всех ветвях процесса на рис. 1, что невозможно без разработки имитационной СС, соответствующей методике проведения СИМ по МДМ в рамках поставленной задачи [2, 3].

Конкретика задачи помогает ЛПР определить смысл вероятностей p_F и p_G , которые представляют собой случайные числовые величины, распределенные в пределах $[p_{\text{MIN}}; p_{\text{MAX}}]$ и отражающие суждения экспертов, сформированные ими на основе как совокупности объективных предпосылок (расчетов, логических заключений, теоретических соображений), так и под влиянием субъективных факторов (предпочтения, личный опыт, внешние воздействия). Если пределы $[p_{\text{MIN}}; p_{\text{MAX}}]$ для каждой случайной числовой величины экспертами определены, значения p_F и p_G могут быть найдены с помощью технологии ММК или в соответствии с "принципом безразличия" приняты соответствующими равновероятным событиями.

Отметим также, во-первых, что каждая ветвь процесса на рис. 1 предполагает отдельный (независимый от других) путь реализации проекта, поэтому общим для них является конечный выигрыш F_N в составе ФОП. Все другие параметры и показатели являются индивидуальными, что, вообще говоря, должно быть отражено их двойными и тройными индексами — не будем делать этого для простоты записи формул, которая не влияет на ход анализа и достоверность его результатов.

Во-вторых, применение правила дизъюнкции "ИЛИ" на схеме рис. 1 не соответствует ни разделению вероятностей p_F и p_G , относящихся к полученным путем разветвления на k -м шаге процесса двум и более смежным вариантам его реализации, ни, тем более, аналогичному разложению полезностей Q_k . Значения p_F определяются субъективным желанием ЛПР продолжить проект на k -м шаге для достижения промежуточного результата (выигрыша или проигрыша) F_k — они, безусловно, связаны с предыдущими действиями ЛПР, но формализовать эту связь настолько трудно, что проще

считать этот выбор каждый раз самостоятельным и независимым от предыстории развития процесса. Точно так же значения p_G определяются суждением ЛПР о наличии или отсутствии корпоративных ресурсов для каждого полученного после разветвления процесса варианта его реализации. Однако если моделируемые случайные события составляют полные группы, суммы вероятностей p_F и p_G в них должны равняться единице, иначе идентичность и симметрия анализа будут нарушены, что недопустимо. Границы случайной числовой величины и детерминированные компоненты ФОП в составе (1) определяются как общей логикой развития процесса, так и тактикой реализации его отдельных (условно самостоятельных) вариантов. Для определения их значений используется функционально-стоимостной анализ, который при проектировании СОИБ предусматривает перечни промежуточных целей и ресурсов, необходимых для их достижения.

В-третьих, на втором этапе, когда процесс на рис. 1 является сходящимся, выполнение проекта можно сравнить с моделированием строительства многоэтажного здания. Поскольку модель фундамента здания предусматривает разные варианты его реализации, а строительство первого этажа без фундамента невозможно, варианты создания фундамента целесообразно привести к общему промежуточному финишу на K -м уровне (см. рис. 1), после которого начинается моделирование строительства первого этажа и т. д. В связи с этим применение правила конъюнкции "И" (см. уровни K , L и N на рис. 1) также не предполагает суммирование вероятностей p_F и p_G , сложение выигрышей F_k и ресурсов G_k , а тем более объединение полезностей Q_k по аддитивной схеме. Объединяя ветви процесса на обведенном штриховым контуром участке схемы рис. 1, например, мы должны фактически заново рассчитать все компоненты полезности Q_{L3} , используя в качестве исходных данных параметры, входящие в состав Q_{K3} и Q_{K4} .

Техническую сторону СИМ по МДМ, связанную с применением ФОП и функционально-стоимостного анализа, рассмотрим более подробно на примере бизнес-процесса проектирования системы активной защиты конфиденциальной информации (САЗ КИ) от утечки из офиса через случайные антенны разного вида [9, 10] — как одного из важных и "научекоемких" фрагментов СОИБ. Применение термина "бизнес-процесс" при этом оправдано тем, что работа по защите КИ представляет собой важную часть бизнеса корпорации, нацеленную на обеспечение ее рыночной эффективности и конкурентоспособности.

Моделирование бизнес-процесса разработки фрагмента СОИБ

Логическая диаграмма, раскрывающая содержание бизнес-процесса разработки САЗ КИ, представлена на рис. 2. Будем рассматривать апертурные (АСА) и пространственно-распределенные (РСА) случайные антенны, полагая, что электромагнитный канал утечки КИ может быть сформирован тремя основными способами:

- по эфиру через совокупность АСА в виде окон и дверей;

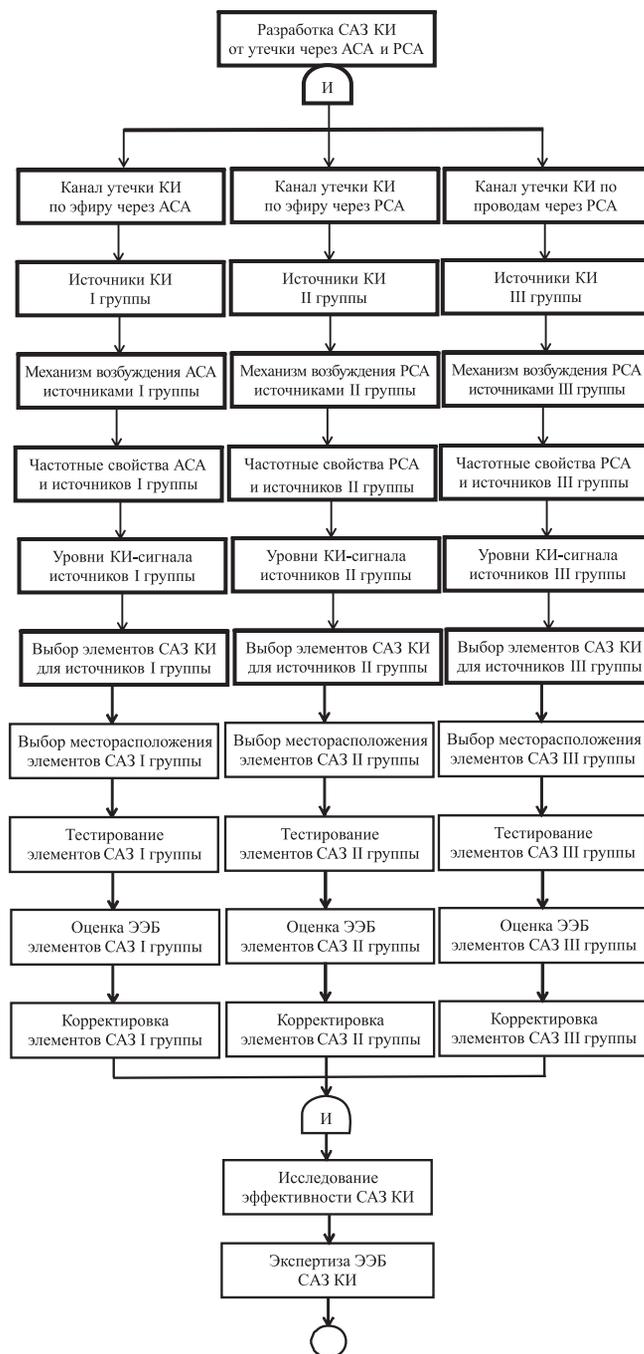


Рис. 2. Логическая диаграмма, раскрывающая процесс разработки САЗ КИ от утечки во внешнюю среду через АСА и РСА: ЭЭБ — эколого-эргономическая безопасность

- по проводам через РСА в виде совокупности линий связи, элементов сетей электропитания, управления, заземления;
- по эфиру через РСА в виде ее перечисленных проводных частей плюс несущие металлические элементы конструкции здания (опоры, балки, решетки и т. п.).

Во втором случае злоумышленник (хакер, недобросовестный конкурент) свои средства перехвата КИ должен подключать к РСА непосредственным образом: последовательно в разрывы проводов или параллельно между проводами. В первом и третьем случаях он может использовать дистанционные способы съема КИ, без нарушения целостности конструкции РСА. Элемент конъюнкции "И" в верхней части схемы означает необходимость учета всех трех вариантов формирования канала утечки КИ одновременно. Выделим в соответствии с номерами указанных вариантов I, II и III группы источников, возбуждающих каналы утечки КИ и отметим, что реальные источники (элементы компьютерных сетей и СС связи и др.) могут принадлежать разным группам одновременно.

Далее определим в каждой группе механизм возбуждения конкретных АСА и РСА источниками КИ и сопоставим пространственно-частотные области, соответствующие им, поскольку для эффективного функционирования канала утечки необходимо, чтобы эти области у случайных антенн и источников максимально совпадали — для нас это самый худший случай.

Следующий "шаг" (прецедент бизнес-процесса) связан с прогнозированием (если определение другим путем невозможно) уровней КИ-сигнала, возбуждающего АСА и циркулирующего в РСА для всех источников I, II и III групп. Оценив эти уровни, ЛПР может обоснованно выбрать элементы САЗ, необходимые для защиты КИ от утечки по каналам, соответствующим источникам I, II и III групп. Указанный выбор проводится, во-первых, по действующим нормам на отношение "помеха/сигнал" в заданной пространственно-частотной области или с учетом других критериев (см. последующий пример). Во-вторых, с применением ФСА при этом оценивается соответствие проекта САЗ объему выделенных ресурсов (выполнение условия ограничения по затратам), проще говоря, составляется и проверяется смета расходов.

Применение ФОП при проектировании СОИБ

Зафиксируем данный момент в процессе проектирования САЗ КИ, по-

скольку теперь есть возможность рассмотреть конкретный пример определения ФОП для расходящегося, а затем сходящегося подпроцесса (см. рис. 1), интересный тем, что значения F_k и G_k не фигурируют здесь в денежном виде. Для источников I группы (см. выделенную верхнюю часть схемы на рис. 2) структура электромагнитного поля сигнала, создаваемого АСА на площади зоны возможного расположения средств перехвата КИ [16], имеет вид, показанный на рис. 3. Задавая варианты расположения источников КИ в здании, разработчик САЗ определяет уровни КИ-сигнала в децибелах, соответствующие G_k , а затем, аналогичным образом, — уровни преднамеренных помех, необходимых для активной защиты КИ, соответствующие F_k . Отношение "помеха/сигнал", от которого зависит эффективность САЗ, при этом будет соответствовать разности $F_k - G_k$, которая фигурирует в формуле (1). Вероятность p_G , определяемая ЛПР-экспертами, при этом отражает случайный характер возникновения утечки КИ-сигнала на объекте защиты; вероятность p_F — наличие в составе САЗ необходимых источников помех; вероятность p_k — возможность перехвата КИ, т. е. все элементы ФОП здесь имеют очевидный смысл и участвуют в моделировании САЗ.

Поскольку число частот, на которых необходимо оценить предполагаемую эффективность САЗ (по аналогии с частотами, указанными на рис. 3), достаточно велико [9, 10], рассматриваемый подпроцесс на первоначальной стадии является рас-

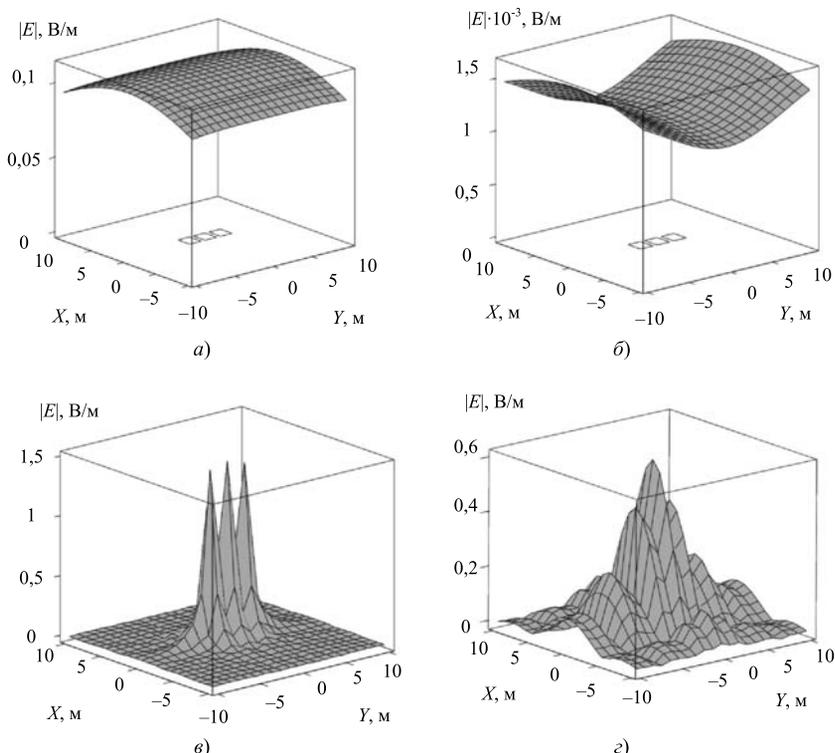


Рис. 3. Распределения модуля напряженности электрического поля АСА на частоте: а — 10 кГц; б — 1 МГц; в — 500 МГц; з — 1 ГГц

ходящимся. Однако затем, когда ЛПР необходимо выбрать лучший, по его мнению, вариант реализации САЗ, подпроцесс очевидным образом сходится. Хотя весь бизнес-процесс создания САЗ КИ (см. рис. 2) на этом не заканчивается: его заключительная часть (реализация и внедрение САЗ) предусматривает окончательный выбор мест установки элементов САЗ в выделенных помещениях здания для источников I, II и III групп. Далее расставленные в здании элементы САЗ всех трех групп проверяются и тестируются на предмет сохранения работоспособности в реальных условиях (которые могут существенно отличаться от заводских и лабораторных), после чего осуществляется оценка их эколого-эргономической безопасности (ЭЭБ).

Если на рабочих местах персонала требования ЭЭБ не выполняются, проводится корректировка условий эксплуатации элементов САЗ поочередно для I, II и III групп: либо меняется их месторасположение, либо рабочие места переносятся в безопасные зоны внутри (за пределами) выделенных помещений. После этого монтаж САЗ КИ считается законченным и начинается ее экспертиза в целом: элемент конъюнкции "И" в нижней части схемы рис. 2 означает дальнейшее рассмотрение всех элементов САЗ для I, II и III групп совместно и одновременно.

Отметим, что данная часть бизнес-процесса "встроена" в надсистему работ (проверки, обследования и исследования) по выявлению технических каналов утечки КИ, методы и порядок проведения которых, применительно к информационной защите случайных антенн, освещены в работе [9]. Оценку эффективности корпоративной САЗ КИ с учетом ЭЭБ для окружающей среды содержат протоколы испытаний и предписание на ее эксплуатацию.

Заключение

Конвергенция принципов и критериев теории риска и теорий ожидаемой полезности в сочетании с возможностями СИМ по МДМ и технологии ММК позволяет повысить эффективность управления СС организационно-технического типа, что показано на примере разработки фрагмента СОИБ корпорации в виде САЗ КИ. Представлены схема определения ФОР, сочетающего признаки критериев объективной и субъективной ожидаемой полезности, а также логическая диаграмма, раскрывающая процесс разработки САЗ КИ от утечки во внешнюю среду через АСА и РСА. При

формировании ФОР, ввиду его универсальности, могут быть использованы как стоимостные, так и не связанные с денежной оценкой показатели. Бизнес-процесс разработки САЗ КИ является частью работ по созданию СОИБ и управлению безопасностью корпорации в интересах обеспечения ее рыночной эффективности и конкурентоспособности.

Список литературы

1. **Нейман Дж. фон, Моргенштерн О.** Теория игр и экономическое поведение: Пер. с англ. М.: Наука, 1970. 708 с.
2. **Димов Э. М., Маслов О. Н., Пчеляков С. Н., Скворцов А. Б.** Новые информационные технологии: подготовка кадров и обучение персонала. Часть 2. Имитационное моделирование и управление бизнес-процессами в инфокоммуникациях. Самара: Изд-во СНЦ РАН, 2008. 350 с.
3. **Димов Э. М., Маслов О. Н., Трошин Ю. В.** Снижение неопределенности выбора управленческих решений с помощью метода статистического имитационного моделирования // Информационные технологии. 2014. № 6. С. 51–57.
4. **Королев В. Ю., Бенинг В. Е., Шоргин С. Я.** Математические основы теории риска. М.: Физматлит, 2007. 544 с.
5. **Маслов О. Н.** Моделирование неопределенностей // Нейрокомпьютеры: разработка, применение. 2014. № 9. С. 79–84.
6. **Винер Н.** Творец и робот: Пер. с англ. М.: Прогресс, 1996. 104 с.
7. **Методы комплексного контроля безопасности информации на объектах телекоммуникационных систем органов государственного управления.** М.: Изд. УДП РФ, 2009. 368 с.
8. **Маслов О. Н.** Безопасность корпорации: моделирование и прогнозирование внутренних угроз методом риска. Самара: Изд-во ПГУТИ—АЭРОПРИНТ, 2013. 170 с.
9. **Маслов О. Н.** Случайные антенны: теория и практика. Самара: Изд-во ПГУТИ—ОФОРТ, 2013. 480 с. URL: http://eis.psuti.ru/images/books/sluch_ant
10. **Маслов О. Н.** Применение метода статистического имитационного моделирования для исследования случайных антенн и проектирования систем активной защиты информации // Успехи современной радиоэлектроники. 2011. № 6. С. 42–55.
11. **Основы теории управления в системах специального назначения.** М.: Изд. УДП РФ, 2008. 400 с.
12. **Виттих В. А.** Организация сложных систем. Самара: Изд-во СНЦ РАН, 2010. 66 с.
13. **Голенко Д. И., Димов Э. М., Осокин В. В.** Вопросы оптимального планирования в системах управления дискретно-непрерывными производством. Опыт применения прикладных методов математики и вычислительной техники в народном хозяйстве. М.: Статистика, 1974. С. 15–32.
14. **Savage L. J.** The Foundations of Statistics. N. Y.: Wiley, 1954. 310 p.
15. **Маслов О. Н., Фролова М. А.** Функционал ожидаемой полезности: принципы моделирования и практического применения // Инфокоммуникационные технологии. 2015. Т. 13, № 3. С. 291–297.
16. **Маслов О. Н., Раков А. С., Силкин А. А.** Статистические характеристики поля решетки апертурных случайных антенн // Радиотехника и электроника. 2013. Т. 58, № 11. С. 1093–1101.

Information Security Management Corporation of Corporation Using the Risk Criterion and Expected Utility

The problem of convergence principles and quantitative criteria of theories of knowledge, risk and expected utility are considered in the article. Convergence allows the use criteria for risk acceptance and the expected utility in the statistical simulation modeling (SSM) by Dimov—Maslov method (MDM). The purpose of SSM by MDM is the study and management of complex systems (SS) organizational and technical type. An example of the SS organization and technical type is the system of information security (SIS) of the corporation. The main problem of modeling is to reduce the impact of uncertainties of initial and intermediate data on the SSM results by MDM. And this problem is solved with the help of computer technology of Monte Carlo method. The effectiveness of reducing uncertainty of the administrative decisions choice at SIS designing. Initial data for SSM by MDM formed on the basis of verified and axiological knowledge about the object. The criterion of justified risk, which corresponds to the expected utility functional is formulated to select management decisions. The diagram of formation of expected utility functional during SSM by MDM. The diagram of formation of expected utility functional during SSM by MDM is presented in this article. Conceptual and logical diagram of a business process of development active protection systems of confidential information is presented in the article. Ways of analyzing the effectiveness of SIS corporations using SSM by MDM are planned.

Keywords: corporate management, information security, statistical simulation modeling, risk, expected utility

References

1. **Nejman Dzh. fon, Morgenshtern O.** *Teorija igr i jekonomicheskoje povedenie*. Per. s angl. Moscow: Nauka, 1970. 708 p.
2. **Dimov Je. M., Maslov O. N., Pcheljakov S. N., Skvorcov A. B.** *Novye informacionnye tehnologii: podgotovka kadrov i obuchenie personala. Chast' 2. Imitacionnoje modelirovanie i upravlenie biznes-procesami v infokommunikacijah*. Samara: Izd-vo SNC RAN, 2008. 350 p.
3. **Dimov Je. M., Maslov O. N., Troshin Ju. V.** Snizhenie neopredelennosti vybora upravlencheskih reshenij s pomoshh'ju metoda statisticheskogo imitacionnogo modelirovanija, *Informacionnye tehnologii*, 2014, no. 6, pp. 51—57.
4. **Korolev V. Ju., Bening V. E., Shorgin S. Ja.** *Matematicheskie osnovy teorii riska*. Moscow: Fizmatlit, 2007. 544 p.
5. **Maslov O. N.** Modelirovanie neopredelennostej, *Nejrokompjutery: razrabotka, primenenie*, 2014, no. 9, pp. 79—84.
6. **Viner N.** *Tvorec i robot*. Per. s angl. Moscow: Progress, 1996. 104 p.
7. **Metody kompleksnogo kontrolja bezopasnosti informacii na objektah tele-kommunikacionnyh sistem organov gosudarstvennogo upravlenija. Moscow: Izd. UDP RF, 2009. 368 p.**
8. **Maslov O. N.** *Bezopasnost' korporacii: modelirovanie i prognozirovanie vnutrennih ugroz metodom riska*. Samara: Izd-vo PGUTI—AJeROPRINT, 2013. 170 p.
9. **Maslov O. N.** *Sluchajnye anteny: teorija i praktika*. Samara: Izd-vo PGUTI—OFORT, 2013. 480 p. URL: http://eis.psuti.ru/images/books/sluch_ant
10. **Maslov O. N.** Primenenie metoda statisticheskogo imitacionnogo modelirovanija dlja issledovanija sluchajnyh antenn i proektirovanija sistem aktivnoj zashhity informacii, *Uspehi sovremennoj radioelektroniki*, 2011, no. 6, pp. 42—55.
11. **Osnovy teorii upravlenija v sistemah special'nogo naznachenija. Moscow: Izd. UDP RF, 2008. 400 p.**
12. **Vittih V. A.** *Organizacija slozhnyh sistem*. Samara: Izd-vo SNC RAN, 2010. 66 p.
13. **Golenko D. I., Dimov Je. M., Osokin V. V.** Voprosy optimal'nogo planirovanija v sistemah upravlenija diskretno-nepreryvny-mi proizvodstvom, *Opyt primenenija prikladnyh metodov matematiki i vychislitel'noj tehniki v narodnom hozjajstve*. Moscow: Statistika, 1974, pp. 15—32.
14. **Savage L. J.** *The Foundations of Statistics*. N. Y.: Wiley, 1954. 310 p.
15. **Maslov O. N., Frolova M. A.** Funkcional ozhidaemoj poleznosti: principy modelirovanija i prakticheskogo primenenija, *Infokommunikacionnye tehnologii*, 2015, vol. 13, no. 3, pp. 291—297.
16. **Maslov O. N., Rakov A. S., Silkin A. A.** Statisticheskie harakteristiki polja reshetki aperturnykh sluchajnyh antenn, *Radiotekhnika i jelektronika*, 2013, vol. 58, no. 11, pp. 1093—1101.

А. В. Еременко¹, канд. техн. наук, доцент каф.

"Инфокоммуникационные системы и информационная безопасность", e-mail: nexus@mail.ru,

А. Е. Сулавко², канд. техн. наук, ст. преподаватель каф.

"Комплексная защита информации", e-mail: sulavich@mail.ru,

Е. В. Толкачева¹, канд. техн. наук, доцент каф. "Инфокоммуникационные системы и информационная безопасность", e-mail: tolkacheva_ev@mail.ru,

Е. А. Левитская³, инженер-исследователь, e-mail: laska_kb@mail.ru

¹ Омский государственный университет путей сообщения (ОмГУПС (ОмИИТ)), г. Омск,

² Омский государственный технический университет (ОмГТУ), г. Омск,

³ ФГУП "Российский Федеральный Ядерный Центр — Всероссийский научно-исследовательский институт технической физики имени академика Е. И. Забабахина", г. Снежинск

Метод защиты текстовых документов на электронных и бумажных носителях на основе скрытого биометрического идентификатора субъекта, получаемого из подписи¹

Рассмотрена проблема защиты авторского права (интеллектуальной собственности), возникающего при создании текстового произведения. Объектом исследования в работе выступают методы кодирования информации в текстовых контейнерах. Предложен метод встраивания в текстовый документ цифрового водяного знака, основанного на биометрических признаках автора документа. Определена информационная емкость разработанного метода. Разработан способ проверки документов на электронных и бумажных носителях на предмет неправомерного изменения и их аутентичности.

Ключевые слова: биометрические признаки, защита интеллектуальной собственности, помехоустойчивое кодирование, цифровой водяной знак, стеганография, ключевая последовательность

Введение

Развитие информационных сетей постепенно превращает мир в единое информационное и коммуникационное пространство, обостряя проблемы, связанные с защитой авторских прав. Помимо технической стороны вопроса защиты интеллектуальной собственности, актуальным остается вопрос законодательства. В настоящее время действие закона № 187-ФЗ распространяется только на фильмы, но от "пиратства" страдают и другие творческие деятели. По мнению генерального директора ЭКСМО, прозвучавшему на конференции "Право на Download 2013", за минувший год потери отрасли книгоиздания от интернет-пиратства составили 7,5 млрд руб. Согласно прогнозу к 2015 г. ущерб составит уже 30 млрд руб., а к 2018 г. достигнет значения 70 млрд руб. [1]. Основным подходом к техническому решению проблемы нелегального копирования является добавление в каждую распространяемую копию скрытого идентификатора (водяного знака), позволяющего однозначно определить источник утечки. Технологии встраивания цифрового идентификатора разрабатывали для электронных документов. Однако при печати текстовых документов возникает проблема переноса скрытого идентификатора автора на бумагу. Цель работы —

определить потенциал подписи субъекта для создания на ее основе скрытого биометрического идентификатора и разработать метод встраивания идентификатора в текстовые документы на электронных и бумажных носителях для подтверждения их целостности и аутентичности.

1. Стеганографические методы защиты от нелегального копирования авторских произведений

На сегодняшний день существует большое число систем внедрения цифровых водяных знаков (ЦВЗ) в мультимедийную информацию. Менее проработанным является вопрос защиты текстовой информации с помощью внедрения ЦВЗ. В лингвистической стеганографии выделяют следующие направления: использование особенностей символов, кодирование смещением строк, кодирование смещением слов, синтаксические и семантические методы. Их достоинства и недостатки рассмотрены в работе [2]. В литературе [3, 4] также можно встретить описание синтаксических и семантических методов внедрения информации, однако отсутствует описание их адаптации для внедрения ЦВЗ.

Настоящая работа посвящена разработке технического метода защиты авторского права (интеллектуальной собственности), возникающего при создании текстового произведения. Исходя из постановки задачи, содержание текста после встраива-

¹ Работа выполнена при финансовой поддержке РФФИ (грант № 15-37-50366).

ния защитной метки не должно изменяться, так как должен быть сохранен авторский стиль произведения, в котором одинаковое значение имеют пунктуационные знаки и слова. Данное условие и требование сохранения возможности проверки защитной метки после вывода документа на печать накладывают ограничения на выбор стеганографического метода встраивания ЦВЗ. Аналитическое исследование проблемы показало, что только кодирование вертикальным смещением строк и горизонтальным смещением слов позволяет решить поставленную задачу, так как при таком подходе изменяется оформление документа, а не его содержательная часть. С учетом опыта [4] необходимо проработать вопрос генерации защитной метки и ее встраивания в различные участки текстового документа для обеспечения связи содержимого документа и идентификационных признаков подписи автора произведения.

В соответствии с предлагаемой технологией при завершении работы над текстом автор произведения (текстового документа) его подписывает. В качестве подписи может быть использован не только автограф, но и текстовый пароль. Из образа подписи автора формируется криптографический ключ. Далее на основе хеша текста и ключа генерируется биометрический идентификатор (защитная метка), который с помощью методов стеганографии связывается непосредственно с содержимым документа с возможностью его восстановления для проверки авторства и целостности текста после печати.

2. Информативные признаки подписи авторов текстовых произведений

Число признаков, необходимое для успешного решения некоторой задачи генерации ключа, зависит от их информативности. В рамках работы по гранту РФФИ 15-07-09053 "Исследование методов получения криптографических ключей шифрования из динамических биометрических характеристик пользователей компьютерных систем" авторами настоящей работы были апробированы различные виды ортогональных базисных функций для описания динамики подписи и оценена информативность признаков, получаемых на их основе.

Для создания эталона подписи пользователь несколько раз воспроизводит ее на графическом планшете. Из каждой реализации подписи вычисляются параметры, характеризующие внешний вид подписи — расстояния между координатами контурного образа, отношение длины к ширине, центр подписи, угол наклона, и ее динамику — коэффициенты корреляции между функциями динамики подписи, коэффициенты вейвлет-преобразования. После этого выполняется расчет средних значений указанных параметров по всем реализациям. Полученный вектор значений является эталоном подписи. Затем эталонные значения признаков округляют и представляют в виде последовательности m бит A_m . Описание процедуры округления будет представлено ниже.

3. Встраивание защитной метки в текст документа

Метод встраивания защитной метки в документ с помощью вертикального сдвига строк и горизонтального смещения слов описывается в работе [4]. В работе рассматриваются различные виды искажений в изображении документа (шум, перекося строк, растяжение и сжатие текста, размытость, случайный сдвиг и равномерное изменение интенсивности), возникающие при его ксерокопировании, сканировании и выводе на печать, и предложены способы компенсации данных искажений.

Изображение страницы может быть представлено в виде функции $f(x, y) = 0$ или 1 , $x \in [0, W]$, $y \in [0, L]$, где W и L — ширина и длина документа, зависящие от разрешения сканирующего устройства. Для простоты примем, что $x \in R$, $y \in R$ и функция $f(x, y)$ принимает непрерывные значения. Изображение текстовой строки описывается функцией $f(x, y) = 0$ или 1 , $x \in [0, W]$, $y \in [t, b]$, где t и b — верхняя и нижняя границы строки.

Горизонтальное представление строки на изображении можно описать интегральной величиной $h(y)$, т. е. суммой всех интервалов, где функция $f(x, y)$

принимает ненулевые значения $h(y) = \int_0^W f(x, y) dx$,

$y \in [t, b]$. Вертикальная область текстовой строки может быть представлена суммой интервалов в вертикальном направлении, где функция $f(x, y)$ при-

нимает ненулевые значения $v(x) = \int_t^b f(x, y) dy$, $x \in [0, W]$.

Для снижения ошибок декодирования при возникновении искажений в изображении документа предлагается кодировать строку только в том случае, если она и ближайшие соседние строки полностью заполнены. Один бит сообщения может быть закодирован в тексте через смещение строки вверх или вниз по вертикали от стандартной позиции, определяемой межстрочным интервалом, заданным оформителем документа. При этом соседние строки предлагается использовать как контрольные и их положение не должно изменяться. Также для кодирования битов можно использовать сдвиг слова по горизонтали влево или вправо. Для этого в строке выделяют три группы слов: два слова по краям строки и оставшиеся слова между ними, объединенные в одну среднюю группу. Слова по краям строки рассматривают как контрольные, и их положение не должно изменяться по завершении операции кодирования. Таким образом, для описания процедуры кодирования можно оперировать единой моделью представления текста в виде блоков (рис. 1): средний блок используется для ко-

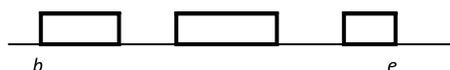


Рис. 1. Модель представления текста в виде блоков, используемых для кодирования и декодирования битов сообщения

дирования бита сообщения, крайние блоки являются контрольными и их положение не изменяется по завершении операции кодирования.

Принципы кодирования защитной метки. Процедура кодирования битов защитной метки в документе выглядит следующим образом.

1. В отличие от работы [4], где для кодирования битов сообщения предложено использовать каждую вторую строку, соответствующую указанным выше условиям, в настоящей работе в целях увеличения емкости метода для кодирования выбирают строки, составляющие абзац, который является структурной единицей при форматировании документа (изменение межстрочного интервала применяют ко всем строкам, входящим в абзац). Положение первой и последней строк абзаца остается неизменным и используется для вычисления оригинального межстрочного интервала. При этом смещение выбранных для кодирования строк необходимо проводить в одном направлении применительно ко всему документу (либо вверх, либо вниз).

2. В отличие от работы [4], где кодирование бита сообщения предложено выполнять смещением центральной группы слов по горизонтали, в целях увеличения емкости метода в настоящей работе предлагается использовать все слова в строке за исключением первого и последнего, которые являются вспомогательными при определении оригинального расстояния между словами (до изменения их положения в строке). Смещение выбранных для кодирования слов также необходимо выполнять в одном направлении (вправо или влево). Смещение кодирует 1, отсутствие смещения — 0.

3. Защитную метку встраивают в каждую страницу документа. В зависимости от наполненности страницы текстом (емкости контейнера) проводят формирование защитной метки с числом бит 2^n с соблюдением условия $2^n \leq E$, где E — доступная для встраивания сообщения емкость контейнера в битах. При этом заполняется весь доступный объем контейнера. Зависимость надежности проверки защитной метки от числа n будет оценена в другом разделе настоящей работы.

4. Операция смещения строки является приоритетной перед операцией сдвига слова при кодировании очередного бита сообщения.

Декодирование (считывание) защитной метки из цифровой копии документа. После получения изображения документа в электронном виде проводим декодирование защитной метки. Рассмотрим обобщенную модель, которая описывает оба способа кодирования бит сообщения. Модель представляет собой область $h(x)$, $x \in [b, e]$ (см. рис. 1), состоящую из трех блоков. В случае если кодирование проводим смещением слов, область $h(x)$ состоит из трех блоков слов, первый и последний из которых являются контрольными и их положение не изменяется. Если речь идет о модели кодирования сдвигом строк, то $h(x)$ обозначает область, состоящую из трех

блоков строк, первый и последний из которых являются контрольными и не изменяют своего положения по завершении операции кодирования.

Декодирование строк. Для определения смещения строк выполняют расчет центра тяжести каждого из трех блоков, входящих в область $h(x)$, определяемых соответственно тремя интервалами $[b_1, e_1]$, $[b_2, e_2]$ и $[b_3, e_3]$.

Центр тяжести блока определяем по формуле

$$c_i = \frac{\int_{b_i}^{e_i} xh(x)dx}{\int_{b_i}^{e_i} h(x)dx},$$

где i — номер блока, $i = 1, 2, 3$.

Достаточно определить ординату центра тяжести каждого блока для проверки факта смещения строки. На следующем шаге по ординатам центров тяжести y_1 , y_2 и y_3 определяем расстояние между блоками. Признаком того, что строка сдвинута вверх, является выполнение неравенства $|y_2 - y_1| < |(y_3 - y_1)/2|$. Выполнение неравенства $|y_2 - y_1| > |(y_3 - y_1)/2|$ означает, что строка смещена вниз.

Декодирование слов. Пусть мы имеем область $h(x)$ и три интервала $[b_1, e_1]$, $[b_2, e_2]$ и $[b_3, e_3]$, которые определяют три группы слов. Полагаем, что $h(x) = 0$ в интервалах между группами слов. Среднюю группу используем для кодирования сообщения в тексте, в то время как положения крайних слов остаются неизменными.

Определим область, описывающую строку со сдвигом центрального блока влево, как

$$h'(x) = \begin{cases} h(x), & x < b_2 - \varepsilon \text{ или } x > e_2; \\ h(x + s), & b_2 - \varepsilon \leq x \leq e_2 - \varepsilon; \\ 0, & e_2 - \varepsilon \leq x < e_2. \end{cases}$$

После сдвига центрального блока вправо $h'(x)$ описывает новую область, получившуюся из $h(x)$:

$$h''(x) = \begin{cases} h(x), & x < b_2 \text{ или } x > e_2 + \varepsilon; \\ 0, & b_2 \leq x < b_2 + \varepsilon; \\ h(x - s), & b_2 + \varepsilon \leq x \leq e_2 + \varepsilon, \end{cases}$$

где ε — значение сдвига, не превышающее половины расстояния между словами.

Обозначим $h'(x)$ область строки, полученную после оцифровывания документа. Для того чтобы определить направление смещения средней группы слов вычислим сумму:

$$\sum_{b_1}^{e_3} h'(x)(h'(x) - h''(x)) \geq 0.$$

Положительный либо равный нулю результат означает, что средняя группа слов была сдвинута влево, т.е. область $h'(x)$ в большей степени совпала с описанием $h'(x)$, и сдвинута вправо в противном

случае. Независимо от примененного типа выравнивания абзаца можно определить положение каждого слова в строке относительно первого и последнего слов. На следующем шаге определяется фактическое положение слов в строке, после чего происходит декодирование содержимого.

Оценка информационной емкости метода и контейнера. Одним из вопросов, который необходимо исследовать в настоящей работе, является определение зависимости сложности криптографического ключа (его надежности) и емкости текста, в который данный ключ может быть встроено.

При разработке системы защиты текстовых документов на электронных и бумажных носителях необходимо учитывать емкость контейнера. Поэтому первым этапом в процессе стеганографии является выбор файла, который необходимо скрыть. Его еще называют информационным файлом. В сопроводительной документации к большинству известных программ по стеганографии говорится, что для сокрытия информации объем памяти файла-контейнера должен примерно в 8 раз превышать объем памяти информационного файла. Следовательно, чтобы спрятать файл размером 710 Кбайт, понадобится графический файл объемом 5680 Кбайт. Но если рассмотреть методы сокрытия информации в тексте, то становится очевидным, что объем памяти файла-контейнера должен в 50—200 раз превышать объем памяти информационного файла [5].

В то же время эффективность методов встраивания дополнительной информации в скрывающие данные (контейнеры) в первую очередь определяется информационной емкостью метода. Приведем несколько интересных оценок лингвистов. Средняя длина слова на корпусе текстов частотного словаря О. Н. Ляшевской и С. А. Шарова составляет 5,28 символа, а на корпусе частотного словаря Л. Н. Засориной — 5,4 символа. Средняя длина предложения в русском языке составляет 10,38 слов [6]. Размер условно-стандартной (учетной) страницы формата А4 равен 1800 знакам с пробелами [7]. В таблице представлен расчет информационной емкости контейнера для страницы формата А4. Нижняя граница оценки получена на основании данных из открытой литературы, верхняя граница — по результатам собственных измерений типовой страницы, плотно заполненной текстом. Число слов в первой строке таблицы получено следующим

Оценка информационной емкости контейнера

Источник данных	Число знаков с пробелами	Число строк	Число слов на странице А4, пригодных для кодирования	Информационная емкость, %
Открытый источник	1800	35	155	1
Собственные измерения	2690	35	267	1,5

образом. Средняя длина слова в русском языке 5,28 символа была округлена в большую сторону до 6. Длина слова с учетом пробелов с обеих сторон составляет 8 символов, тогда число слов в документе равняется 225 (частное от деления 1800 на 8). За вычетом крайних слов в строках ($2 \times 35 = 70$), которые нельзя использовать для кодирования битов сообщения, остается 155 слов, пригодных для внедрения сообщения.

С учетом того, что на каждую строку абзаца и на каждое слово, встреченное в текстовой строке, приходится один бит скрываемого сообщения, средняя информационная емкость предложенного метода невысока. В соотношении количества скрываемой информации к объему контейнера (скрывающего текста) она составляет 1...1,5 %. Применительно к текстовой стеганографии информационная емкость зачастую оценивается не прямым соотношением количества скрываемой информации к исходному объему контейнера, а отношением бит/символ. Другими словами, оценивается среднее число битов, приходящихся на один символ текстового сообщения. Такая оценка продиктована прежде всего тем, что минимальной единицей текстового сообщения является символ, а не отдельный бит. Информационная емкость метода в отношении бит/символ составляет 0,02, что соответствует эффективности использования контейнера на 2 % (1 бит на 48 бит 6-символьного слова). Для сравнения — эффективность использования контейнера методом, основанном на использовании синонимов, при искусственной генерации текстов не превышает 11 % в отношении бит/символ. Примерно ту же эффективность имеют и методы, основанные на использовании пробельных символов и знаков пунктуации.

4. Генерация защитной метки и проверка авторства документа

Чтобы определиться со структурой внедряемой маркировки, необходимо понять, какие требования к ней предъявляются. Для этого следует определить, что включает в себя задача защиты электронного документа. Были выделены следующие части данной задачи:

- защита документа и его частей от подделки;
- проверка авторства документа.

Для решения этих задач было решено внедрять в документы следующую информацию:

- идентификатор автора текстового документа, связанный с биометрическими характеристиками его подписи;
- значение хеш-функции от содержимого документа.

Из этих требований и из описанного выше способа сокрытия информации было принято решение о целесообразности внедрения маркировки в каждую страницу текстового документа. Кроме того, ввиду необходимости проверки наличия изменений в документе, следует привязывать маркировку

не только к ключу, получаемому из биометрических характеристик подписи автора документа, но и к самому контейнеру.

После подготовки произведения автор подписывает документ. Затем необходимо вычислить статические и динамические признаки подписи автора документа, провести оценку информативности каждого признака. Выработку защитной метки проводят постранично, для этого текст с каждой страницы загружают в буфер. С учетом требования распределения данных скрытой маркировки по всему документу, а также необходимости обеспечения возможности проверки маркировки по каждой малой части документа отдельно, алгоритм встраивания маркировки должен включать в себя следующие основные шаги.

1. Провести оценку информационной емкости текстового контейнера описанным ранее способом.

2. С помощью псевдослучайного генератора выработать ключевую последовательность Key , равную m бит. Биометрический идентификатор автора не может быть непосредственно использован для шифрования текстовой строки, так как значения биометрических признаков даже при последовательном вводе нескольких реализаций будут иметь определенный разброс. Ключевую последовательность необходимо кодировать помехоустойчивым кодом, после чего с помощью операции побитового сложения по модулю 2 с биометрическими данными автора получают открытую строку. Ключ после защиты документа уничтожить. Открытую строку, описание произведения и его автора сохранить на специализированном сервере. Связь криптостойкости ключа длиной m бит с надежностью его восстановления из открытой строки, хранящейся на сервере (чем более длинный ключ используется, тем больше малоинформативных признаков применяется для его защиты, следовательно, тем больше ошибки 1-го и 2-го рода при его восстановлении) будет показана ниже.

3. Для получения биометрического идентификатора (защитной метки) используют n -битовую хеш-функцию двух аргументов. Логическая структура рекомендуемой хеш-функции представлена на рис. 2. На вход хеш-функции подать выбранную для маркировки текстовую строку и сгенерированную ключевую последовательность. На первом шаге текстовую строку и ключевую последовательность по-

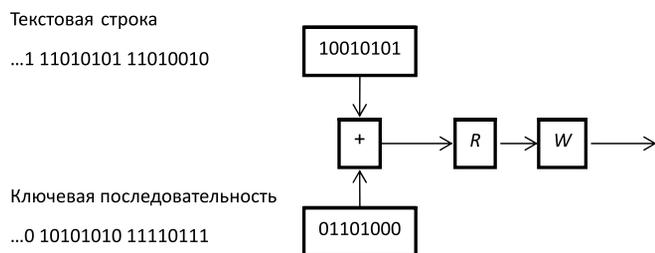


Рис. 2. Вычисление однобитовой хеш-функции от заданных текстовой строки и ключевой последовательности

дают на вход байтового сумматора хеш-функции в байтовом представлении. Байтовый сумматор может быть представлен функцией битового сложения по модулю два (операция XOR) при достаточной длине ключевой последовательности или функцией поточного шифрования текстовой строки на ключе Key . В реализации байтового сумматора с использованием функции битового сложения по модулю два каждый бит байта текстовой последовательности складывается по модулю два с соответствующим битом ключевой последовательности. С выхода байтового сумматора биты последовательно подаются на вход регистра сдвига $R \rightarrow$ с линейной обратной связью. Выбор конкретного примитивного многочлена для регистра сдвига может быть произвольным, в нашем случае выбран многочлен $x^8 + x^4 + x^3 + x^2 + 1$.

Биты с выхода регистра сдвига подаются во временный буфер W , реализующий хранение последних 2^n поступивших в него битов (исходя из условия $2^n \leq E$, где 2^n — длина ключа; E — информационная емкость контейнера). По завершении циклической обработки всех битов последовательности, полученной от исходной текстовой строки с выхода байтового сумматора, последние 2^n битов с выхода регистра сдвига $R \rightarrow$, оставшиеся в буфере W , и формируют результат хеш-функции. Альтернативно можно использовать любую из широко распространенных, проверенных хеш-функций одного аргумента, например функцию SHA-256, подавая ей на вход результат гаммирования двух аргументов.

Для проверки авторства текста субъект, заявляющий данное право, расписывается на графическом планшете. Полученные на основании его подписи биометрические признаки используют для извлечения криптографического ключа из открытой строки (с помощью операции сложения по модулю 2 и последующего применения кода, исправляющего ошибки), хранящейся на сервере. Вычисляют значение хеш-функции от текстовой строки и полученного криптографического ключа способом, описанным выше. Выполняют декодирование защитной метки, содержащейся в тексте и ее сравнение с вычисленной меткой на основании предъявленных данных. При совпадении защитных меток авторство считается подтвержденным, а содержание документа неизменным.

Для определения связи криптостойкости ключа длиной m бит и надежности его восстановления из открытой строки, хранящейся на сервере, был проведен вычислительный эксперимент с использованием предложенного способа восстановления ключа, который заключается в следующем. Предварительно исходные значения признаков кодируют с помощью преобразования вида: $y \in f(x)$, где $x \in X$, X — множество возможных значений признака, $y \in \{0, 1, 3, 7, 15, 31, 63, 127, 255, 254, 252, 248, 240, 224, 192, 128\}$. Выходное значение y представлено в двоичном виде для того, чтобы исклю-

читать неинформативные биты данных. Далее проводят "склеивку" битовых последовательностей в одну результирующую, которую "объединяют" со случайной строкой. В работе [8] показана связь эффективности коррекции ошибок с методами группирования битов с разной вероятностью единичной ошибки. Несмотря на предпринятые в данном направлении усилия, единого подхода для решения этого вопроса до сих пор выработано не было. В настоящей работе предпринята попытка развития данного направления по модернизации нечетких экстракторов. Последовательность признаков задается случайным образом для каждого субъекта индивидуально во время формирования открытой строки. Осуществляется конкатенация битовых представлений только наиболее информативных (стабильных) признаков для субъекта. Для каждого признака по всем отобранным для создания эталона преобразованным реализациям вычисляют относительную частоту появления единичных (или нулевых) бит. Далее определяют интегральную вероятность появления единичного бита во всех разрядах, при этом относительную частоту берут как вероятность. Чем больше разрядов будут иметь частоты, близкие к 0 или 1, тем меньше получится итоговое произведение и тем выше интегральная оценка стабильности (информативности) признака для субъекта. Несложно заметить, что при подсчете частот появления нулевых бит (вместо единичных), значение производящей функции будет тем же. Далее все признаки ранжируют по информативности (изменяется их порядок — от самого информативного к самому малоинформативному) и отбирают определенное число признаков, проводят конкатенацию битовых представлений этих признаков, остальные отбрасывают. Чем больше использовано признаков, тем выше длина ключа, но и выше сумма ошибок 1-го и 2-го рода. Итоговую битовую последовательность используют для формирования открытой строки.

Для кодирования битовой последовательности *Key* в предложенном методе использовали коды, исправляющие ошибки Боуза — Чоудхури — Хоквингема (БЧХ), для декодирования — алгоритм Питерсона — Горенштейна — Цирлера (ПГЦ), по аналогии с [9].

Результат эксперимента показан на рис. 3 и должен быть учтен для определения компромисса между надежностью восстановления ключа и его криптостойкостью. Он составляет: вероятности ошибок выработки идентификатора 1-го рода 0,008...0,064 и 2-го рода 0,03...0,01 при длине криптографического ключа от 64 до 328 бит соответственно. Полученный результат указан для оптимальной исправляющей способности кода (при которой сумма ошибок 1-го и 2-го рода восстановления ключа была наименьшей при заданном числе признаков), вычисляемой в процессе эксперимента. Достоверность указанных на рис. 3 результатов со-

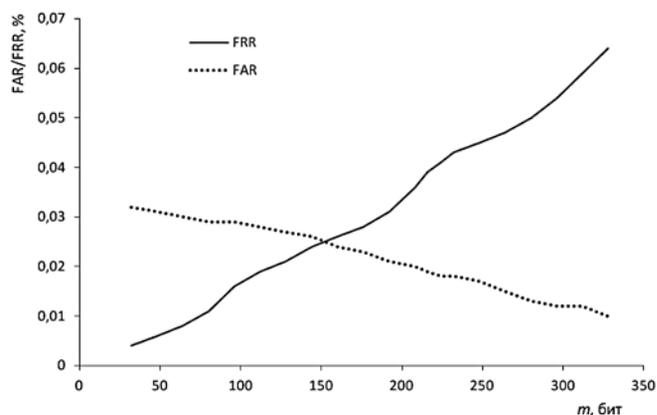


Рис. 3. Связь криптостойкости ключа длиной m бит и надежности его восстановления из открытой строки, хранящейся на сервере

ставляет более 0,98 при доверительном интервале вероятности ошибок 1-го (FRR) и 2-го рода (FAR) 0,01 и 0,002 соответственно, что определялось исходя из экспериментально полученных вероятностей и числа проведенных опытов (8000).

Перед встраиванием защитной метки в текст документа для компенсации возможных ошибок на этапе ее извлечения, связанных с искажением изображения, целесообразно выполнять помехоустойчивое кодирование защитной метки. Следует учитывать, что данная операция увеличит размер защитной метки в несколько раз, что повлечет за собой необходимость использования хеш-функции, возвращающей значение меньшей разрядности ввиду ограничения информационной емкости контейнера. Если объема текста на странице недостаточно для внедрения значения хеш-функции, то он может быть "присоединен" к содержанию предыдущей или последующей страницы.

Заключение

Разработан метод генерации идентификатора автора на основе содержимого документа и биометрических характеристик подписи его владельца с вероятностью ошибок выработки идентификатора 1-го рода от 0,008 до 0,064 и 2-го рода от 0,03 до 0,01 при длине криптографического ключа от 64 до 328 бит соответственно. В качестве способа встраивания защитной метки в документ с учетом поставленной задачи выбрано кодирование битов сообщения смещением слов и строк. Способ обеспечивает скрытность и устойчивость встраиваемого сообщения для определения целостности и аутентичности документа как на электронном, так и на бумажном носителях с соотношением количества скрываемой информации к объему контейнера 1...1,5 %. Информационная емкость предложенного метода достаточна для решения поставленной в работе задачи — защиты документа от незаконного копирования и изменения содержимого. Защитную метку можно внедрять непосредственно в сам до-

кумент, поддерживающий форматирование текста, либо изменения в формате выполнять непосредственно при выводе документа на печать.

Для подавления шумов, возникающих при получении цифровой копии документа, перед выполнением процедуры декодирования сообщения следует применять известные методики, описанные в открытой литературе, например эффективным способом подавления шума типа "salt and pepper" является медианный фильтр [10].

Список литературы

1. **Горвцова М.** Новая "антипиратская" инициатива: окажутся ли под защитой все объекты авторских прав ГАРАНТ.РУ. URL: <http://www.garant.ru/article/501059/#ixzz3PHwgpwoV>. (дата обращения: 15.01.2015).
2. **Балакин А. В., Елисеев А. С., Гуфан А. Ю.** Использование стеганографических методов для защиты текстовой информации // Т-Comm. Спецвыпуск. 2009, апрель. С. 42—50.

3. **Аграновский А. В., Балакин А. В., Хади Р. А.** Запатентованные методы стеганографии в технологиях цифровых водяных знаков // Информационные технологии. 2002. № 9. С. 2—7.
4. **Brassil J., Low S., Maxemchuk N., O'Gorman L.** Document marking and identification using both line and word shifting // Technical report, AT & T Bell Laboratories, 1994. P. 853—860.
5. **Компьютерная стеганография — защита информации или инструмент преступления?** [Электронный ресурс]. URL: <http://www.crime-research.org/library/Steganos.htm>, свободный (дата обращения: 20.12.2011).
6. **Куянов Ю. В., Тришин В. Н.** Количественный анализ Большого русского словаря-справочника синонимов // Научное обозрение: гуманитарные исследования. 2015. № 9. С. 105—111.
7. **Дупленский Н.** Письменный перевод. Рекомендации переводчику и заказчику / под ред. Е. Масловского (СПР). 3 ред. М.: Союз переводчиков России, 2015.
8. **Santos M. F., Aguilar J. F., Garcia J. O.** Cryptographic key generation using handwritten signature // Proc. of SPIE, Orlando, Fla, USA, Apr. 2006. Vol. 6202. P. 225—231.
9. **Еременко А. В., Сулавко А. Е.** Исследование алгоритма генерации криптографических ключей из биометрической информации пользователей компьютерных систем // Информационные технологии. 2013. № 11. С. 47—51.
10. **Сато Ю.** Без паники! Цифровая обработка сигналов. М.: Додэка XXI, 2010. 176 с.

A. V. Eremenko¹, Ph. D., Associate Professor, nexus-@mail.ru,

A. E. Sulavko², Ph. D., Assistant, sulavich@mail.ru,

E. V. Tolkacheva¹, Ph. D., Associate Professor, tolkacheva_ev@mail.ru,

E. A. Levitskaya³, Research Engineer, laska_kb@mail.ru,

¹Omsk State Transport University (OSTU), ²Omsk State Technical University (OmSTU),

³Federal State Unitary Enterprise "Russian Federal Nuclear Center — Academician E. I. Zababakhin All-Russian Research Institute of Technical Physics" (FSUE "RFNC-VNIITF")

The Method of Protection of Electronic and Paper-Based Text Documents with Biometric Identifier of the Subject Obtained from his Signature

The problem of the protection of copyright (intellectual property) that occurs when creating a text works is considered. The object of study in the article are the methods of coding information in the text containers. A method of embedding into a text document a digital watermark based on the biometric features of the author of the document is offered. The informational capacity of this method is determined. A method for inspection of documents in electronic and paper form for unauthorized changes and their authenticity is designed.

Keywords: biometric features, intellectual property protection, error-correcting codes, digital watermark, steganography, the key sequence

References

1. **Gorovcova M.** Novaja "antipiratskaja" iniciativa: okazhutsja li pod zashhitoy vse ob'ekty avtorskih prav GARANT.RU [The new "anti-piracy" Initiative: will be all objects under the protection of copyright GARANT.RU], Access: <http://www.garant.ru/article/501059/#ixzz3PHwgpwoV>. (Date of circulation: 01.15.2015).
2. **Balakin A. V., Eliseev A. S., Gufan A. Ju.** Ispol'zovanie steganograficheskikh metodov dlja zashhity tekstovoj informacii [Using steganography techniques to protect the textual information], *T-Comm*. april 2009, pp. 42—50.
3. **Agranovskiy A. V., Balakin A. V., Hadi R. A.** Zapatentovannye metody steganografii v tehnologijah cifrovых vodjanyh znakov [The patented techniques in the steganographic digital watermark technology], *Information technologies*, 2002, no. 9, pp. 2—7.
4. **Brassil J., Low S., Maxemchuk N., O'Gorman L.** Document marking and identification using both line and word shifting, *Technical report*, AT & T Bell Laboratories, 1994, pp. 853—860.
5. **Комп'ютрная стеганография — зашхита информации или инструмент преступления?** [Computer steganography — information security or

instrument of the crime?]. Access: <http://www.crime-research.org/library/Steganos.htm>, free (date of circulation: 20.12.2011).

6. **Kuyan J. V., Trishin V. N.** Kolichestvennyj analiz Bol'shogo russkogo slovarja-spravochnika sinonimov [Quantitative analysis of large-Russian dictionary of synonyms directory], *Scientific Review: humanities research*, 2015, no. 9, pp. 105—111.

7. **Duplenskiy N.** Pis'mennyj perevod. Rekomendacii perevodchiku i zakazchiku [Translation. Recommendations to the translator and the customer], editor E. Maslowski, 3 red. Moscow, Union of Translators of Russia, 2015.

8. **Santos M. F., Aguilar J. F., Garcia J. O.** Cryptographic key generation using handwritten signature, *Proceedings of SPIE, Orlando, Fla, USA, Apr. 2006*, vol. 6202, pp. 225—231.

9. **Eremenko A. V., Sulavko A. E.** Issledovanie algoritma generacii kriptograficheskikh kljuchej iz biometricheskoj informacii pol'zovatelej komp'juternyh sistem [Investigation of algorithm for generating cryptographic keys from biometric information of users of computer systems], *Informacionnye tehnologii*, 2013, no. 11, pp. 47—51.

10. **Sato Yu.** Bez paniki! Cifrovaja obrabotka signalov [Do not panic! Digital signal processing], Moscow, Dodeka XXI, 2010, 176 p.

А. В. Чернышов, канд. техн. наук, доц., e-mail: sch@mgul.ac.ru,
Московский государственный университет леса

К вопросу о применении оптических дисков для создания долговременных электронных архивных хранилищ информации небольших организаций

Рассмотрена задача минимизации затрат при создании долговременного электронного архивного хранилища информации небольшой организации при одновременном достижении заданного значения надежности хранения информации. Отличительными особенностями такого хранилища являются: использование только оптических дисков однократной записи и неавтоматизированный доступ пользователей к информации на дисках. Построены математические модели вероятности потери части информации и суммарных затрат в архиве без копий и в архиве с заданным числом копий. Выполненные расчеты для разных типов оптических дисков, имеющих разную емкость и стоимость, показывают, что на данный момент оптимально создание хранилища на базе оптических дисков BD-R с выполнением необходимого (в зависимости от заданного значения показателя надежности) числа копий информации.

Ключевые слова: электронное архивное хранилище информации, длительное хранение, оптические носители информации, надежность хранения информации, минимизация затрат

Введение

В последние несколько лет в связи с развитием средств вычислительной техники и, прежде всего, значительным повышением возможностей устройств сканирования, а также устройств хранения информации возрос интерес разных организаций к созданию архивных электронных хранилищ своих документов. Теоретически современные информационные технологии способны избавить организации от необходимости содержать объемные архивы бумажных документов, занимающие немалые площади и требующие определенных условий хранения и обслуживания, путем перевода всех бумажных документов в электронные копии (как минимум, методом простого сканирования, а возможно, и полнотекстовой оцифровки) с последующим их хранением на современных серверах, занимающих относительно небольшие площади.

Однако использование для длительного хранения архивной информации перезаписываемых жестких дисков, повсеместно применяемых в серверах, в ряде случаев вызывает опасение, что сохраняемая на них информация может быть случайно либо умышленно искажена и даже уничтожена. Поэтому современный ГОСТ [1] рекомендует создавать электронные архивы с использованием носителей однократной записи, в качестве которых на сегодняшний день могут рассматриваться только различного типа оптические диски однократной записи (CD-R, DVD \pm R, BD-R). При этом для создания электронных архивов крупных организаций должны быть выполнены несколько условий:

- объем сохраняемой информации исчисляется десятками терабайтов;
- необходим оперативный доступ пользователей к этой информации;

- необходимо минимизировать (в идеале — свести к нулю) вероятность потери даже небольшой части информации;
- необходимо минимизировать эксплуатационные расходы, в том числе энергопотребление, а также затраты на поддержание целостности информации.

Для выполнения этих условий современные производители и поставщики оборудования рекомендуют использовать роботизированные библиотеки оптических дисков [2]. Сразу заметим, что стоимость таких библиотек исчисляется миллионами рублей и, кроме того, ежегодные эксплуатационные расходы для них составляют около 10 % первоначальной стоимости [3]. Позволить себе такие библиотеки могут только действительно крупные организации.

Важно заметить, что бурно развивающиеся в последние несколько лет технологии облачных хранилищ [4] не могут на данный момент считаться надежной технологией длительного хранения архивных данных. Причины этого подробно рассмотрены в [5]. Поэтому для реализации надежного оперативного доступа к архивной информации, да и для обеспечения контролируемого и управляемого ее сохранения необходимо создание и поддержание именно локального электронного архивного хранилища.

На практике требование оперативного доступа к архивной информации может пониматься по-разному, в зависимости от типов работ с информацией, проводимых в организации.

Так, в случаях, характеризующихся необходимостью дистанционного доступа к архивной информации большого числа пользователей, должен быть реализован такой режим доступа, который с точки зрения пользователей ничем бы не отличался от доступа к информации, хранящейся на жестком диске сервера.

Вместе с тем для ряда типов архивных хранилищ частота запросов к документам может быть невысокой. Например, в статье [2] указано, что во многих случаях около 80 % хранимой информации не требует частого обращения, а около 20 % данных никогда не будут запрошены. Так, для многих архивов небольших предприятий и организаций, в частности, школьных музеев, которых на 2010 г. в России существовало более 11 000 [6], число одновременно работающих с архивными данными пользователей исчисляется единицами. Для таких архивов характерно, что их пользователи:

— находятся рядом с архивом, т. е. им не требуется удаленный доступ;

— как следствие, имеют возможность вручную подключить к рабочему компьютеру необходимый носитель данных с архивной информацией (например, вставить нужный оптический диск в привод DVD);

— не ожидают возможности доступа ко всему объему архивной информации только программными методами — без необходимости физического вмешательства в процесс замены оптических дисков в приводе.

Кроме этого, такие архивы, как правило, имеют ограниченное финансирование либо вообще существуют на общественных началах. Однако объемы электронных данных при их оцифровке могут быть значительными и составлять от единиц до десятков терабайтов. Поэтому перед ними также стоит проблема длительного хранения оцифрованных данных в целостной (неискаженной) форме. К тому же для небольших организаций обязательным требованием является минимизация затрат как на аппаратуру, обеспечивающую хранение, так и на обслуживание электронного архива (регулярный контроль целостности данных и др.). И если исходить из утверждения, что именно хранение информации на однократно записываемых оптических дисках дает наилучшие результаты, поскольку такие диски имеют наибольший срок сохраняемости данных (до 50 лет) без необходимости их перезаписи [2, 3], то представляет интерес рассмотреть возможность создания электронного архива отсканированных документов небольшой организации на оптических дисках однократной записи.

Постановка задачи

Как было показано выше, стоимость специализированного оборудования типа роботизированной библиотеки оптических дисков для небольшой организации неоправданно высока, но и потребности в доступе к информации электронного архива, как правило, невелики, поэтому в качестве базового решения рассмотрим простую коллекцию оптических дисков, которые могут храниться в обычном шкафу (сейфе) при соблюдении необходимых условий по освещенности, влажности и температуре. Диски записываются и воспроизводятся по одному в единственном приводе, подключаемом к компьютеру, на котором работает оператор. Установку дисков в привод и их извлечение выполняет сам оператор.

Понятно, что стоимость такого решения складывается из стоимости привода дисков и собственно оптических дисков, число которых будет зависеть от общего объема сохраняемой в архиве информации и емкости одного оптического диска.

Несмотря на то что производителями декларируется высокая надежность сохранения информации на оптических носителях, практика создания электронных архивов показывает, что должно сохраняться несколько копий информации, поскольку в реальной эксплуатации оптические диски со временем все же выходят из строя с некоторой вероятностью [7, 8]. Таким образом, важной составляющей, влияющей и на надежность хранения информации, и на стоимость решения, будет число копий каждого оптического диска.

В результате возникает следующая постановка задачи. Требуется осуществить обоснованный выбор типа оптического носителя информации для создания неавтоматизированного долговременного электронного архивного хранилища информации, обеспечивающего высокую надежность хранения информации (за счет создания необходимого числа копий) при одновременной минимизации затрат.

Однако оценка параметров надежности и стоимости хранения в зависимости от числа сохраняемых копий информации представляет интерес не только для электронного архива небольшой организации, где все манипуляции с дисками выполняются вручную, но и для роботизированного электронного архива крупной организации, поскольку в обоих случаях используются носители информации того же типа.

Архив с минимальной структурой — без резервных копий

Введем следующие обозначения:

W — общий объем данных, сохраняемый в электронном архиве;

V — емкость одного оптического диска;

l — число оптических дисков, необходимых для сохранения информации объема W .

Очевидно, что $l = [W/V]$, где через $[.]$ обозначена операция округления до большего целого.

Структура хранилища электронного архива, не имеющего резервных копий, с учетом введенных обозначений показана на рис. 1.

Стоимость такого архива составит

$$C_1 = cl + C_{\text{п}},$$

где c — стоимость одного оптического диска; $C_{\text{п}}$ — стоимость привода оптических дисков.

Будем считать, что в соответствии с требованиями ГОСТ [1] в архиве используются только однократно записываемые диски. Поэтому какие-либо искажения информации в процессе хранения невозможны, и потеря информации в архиве возможна только за счет того, что какой-либо диск перестал читаться (т. е. вышел из строя). Будем считать, что диск вышел из строя даже в том случае, если с него не удалось прочитать лишь часть информации.

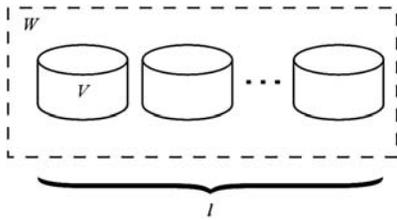


Рис. 1. Структура электронного файлового архива без резервных копий

В соответствии с требованиями ГОСТ контроль исправности архивных дисков (их полной читабельности) должен проводиться с заданной периодичностью, обычно равной одному году. То есть в течение этого периода архив на оптических дисках можно рассматривать как систему без восстановления.

Для оценки надежности сохранения информации применим математический аппарат теории вероятностей [9].

Пусть q — вероятность выхода из строя одного оптического диска в течение периода между контролями исправности. Тогда:

$1 - q$ — вероятность того, что диск в течение этого же периода времени останется исправным (в дальнейшем для краткости упоминание периода времени будем опускать);

$(1 - q)^l$ — вероятность того, что все диски останутся исправными, т. е. будет сохранен весь объем информации W , следовательно, вероятность Q_1 потерять хотя бы часть информации составит

$$Q_1 = 1 - (1 - q)^l.$$

С помощью полученной модели оценим надежность сохранения информации в электронном архиве для различных условий.

Прежде всего зададимся возможными значениями W . В качестве минимального значения выберем 1 Тбайт. Для выбора максимального разумного объема оттолкнемся от емкости предлагаемых сегодня на рынке роботизированных библиотек: Sony ODS-L10 (до 15 Тбайт, стоимость около 1 450 000 руб.) [10] и Panasonic LB-DH8 (в базовой комплектации до 91 Тбайт, стоимость около 5 800 000 руб.) [11]. Для нашего исследования имеет смысл ограничиться объемом 90 Тбайт, имея в виду, что для электронных архивов большинства небольших организаций требуемый объем сохраняемых в архиве данных будет меньше этой величины.

Необходимо также оценить значение величины q . По утверждению производителей оптических дисков типа DVD-R (BD-R), эти диски имеют очень высокую надежность, граничащую с практически гарантированной сохраняемостью информации в течение более 30 лет [2]. Однако имеются исследования, на практических экспериментах показывающие, что порядка 10 % оптических дисков типа CD-R перестают читаться уже через 2 года хранения [7, 8], в то время как производители заявляли о сохраняемости информации на них в течение 100 лет. По дискам типа BD-R на данный момент таких ис-

следований не проводилось (имеющиеся оценки времени сохранения информации основаны на тестах ускоренного старения дисков, что не может дать достоверной картины). Понятно, что на надежность сохранения информации на оптических дисках влияют очень многие факторы, часть из которых не может быть оценена заранее (например, качество изготовления конкретной партии дисков). Поэтому имеет смысл в качестве наихудшего значения задать $q = 0,1$ (т. е. в течение рассматриваемого периода времени из строя выходит 10 % дисков) и в процессе исследования рассмотреть результаты для ряда значений q вплоть до 10^{-10} , имея в виду, что при еще лучших значениях q будут получены и лучшие результаты надежности хранения информации.

Рассмотрим варианты с созданием архивов на базе дисков: CD-R, DVD-R, BD-R, BD-R DL, BD-R XL. Исходные данные, полученные из [12—14], приведены в таблице.

Оценим стоимость одного привода $C_{\text{п}}$ [12]. Используем только максимальное значение стоимости (будем считать, что "выбрали" наилучший по качеству привод), причем для CD-R и DVD-R это будет один и тот же привод стоимостью 2190 руб., а для всех видов DB-R также рассмотрим один привод стоимостью 11 890 руб.

Результаты расчетов для электронного хранилища, имеющего структуру, показанную на рис. 1, представлены на рис. 2—6.

При оценке результатов моделирования будем понимать, что минимальная стоимость одного оптического диска c_{min} на рынке означает, вероятнее всего, что диск имеет низкое качество изготовления и не может быть использован для длительного хранения информации (имеет высокое значение q). Но расчеты с c_{min} позволяют оценить нижнюю границу стоимости хранения информации в архиве. В то же время максимальная стоимость одного оптического диска c_{max} на рынке может косвенно говорить о более высоком качестве диска и большей его пригодности для использования в архиве, однако предсказать его реальную надежность q не представляется возможным. Более того, стоимость одного такого диска будет сильно зависеть от объема партии одновременно закупаемых для архива дисков. Поэтому расчеты с c_{max} позволяют оценить верхнюю границу стоимости решения, но реальная стоимость хранения информации может оказаться немного ниже.

Как можно видеть из рис. 2—6, архив, построенный на базе дисков CD-R, будет иметь неприемлемо высокую стоимость при очень низкой надежности

Тип диска	Емкость диска V , Гбайт	Минимальная цена c_{min} , руб.	Максимальная цена c_{max} , руб.
CD-R	0,7	20	67
DVD-R	4,7	21	106
BD-R	25	65	140
BD-R DL	50	160	510
BD-R XL	100	1100	3606

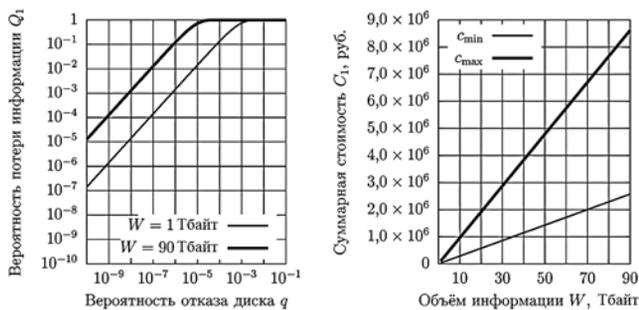


Рис. 2. Результаты расчетов для архива на базе дисков CD-R

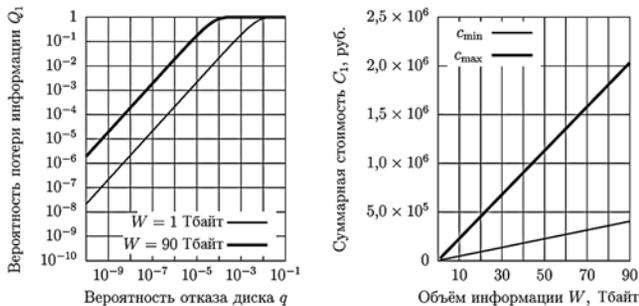


Рис. 3. Результаты расчетов для архива на базе дисков DVD-R

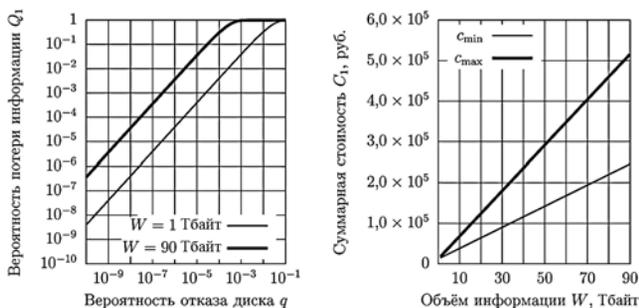


Рис. 4. Результаты расчетов для архива на базе дисков BD-R

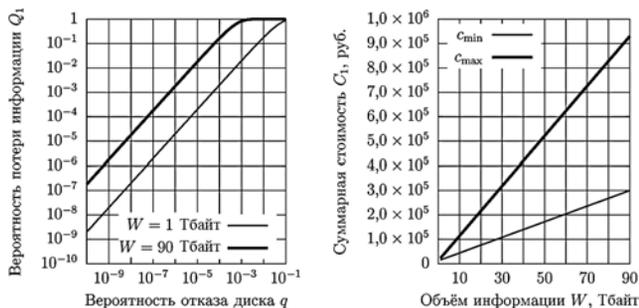


Рис. 5. Результаты расчетов для архива на базе дисков BD-R DL

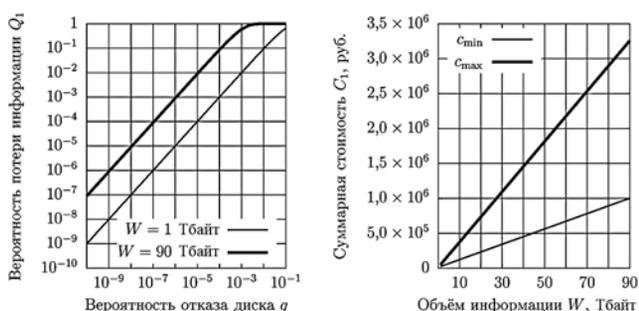


Рис. 6. Результаты расчетов для архива на базе дисков BD-R XL

хранения. Поэтому в дальнейшем такой вариант построения архива рассматриваться не будет.

Все остальные решения дают существенный выигрыш по стоимости относительно предлагаемых на рынке решений Sony [10] и Panasonic [11]. Причем, в порядке увеличения стоимости, решения можно расположить в следующем порядке: BD-R, BD-R DL, DVD-R, DB-R XL. Однако, как и следовало ожидать, надежность хранения информации увеличивается с уменьшением числа дисков (т. е. с ростом их емкости). С этой точки зрения наиболее предпочтительным выглядит решение с применением BD-R XL и наименее предпочтительным — с применением DVD-R.

Впрочем, для рассмотренной структуры электронного архива ни одно из рассмотренных решений не дает удовлетворительного показателя надежности хранения. Для всех вариантов на графиках вероятности потери информации имеются зоны значений q , где величина Q_1 приближается к 1 (а как было показано выше, предсказать заранее значение величины q для конкретной партии дисков невозможно).

Архив с несколькими копиями информации

Для повышения надежности хранения информации в архиве должно быть создано несколько идентичных копий информации. Структура такого архивного хранилища показана на рис. 7.

Для упрощения расчетов можно считать, что в архиве имеется по l копий каждого оптического диска. Тогда:

q^n — вероятность выхода из строя всех копий одного диска (т. е. вероятность потери той части информации, которая записана на этом диске);

$1 - q^n$ — вероятность того, что хотя бы одна копия диска останется целой (информация не будет потеряна);

$(1 - q^n)^l$ — вероятность того, что будет сохранен весь объем информации W , следовательно вероятность Q_n потерять хотя бы часть информации составит

$$Q_n = 1 - (1 - q^n)^l.$$

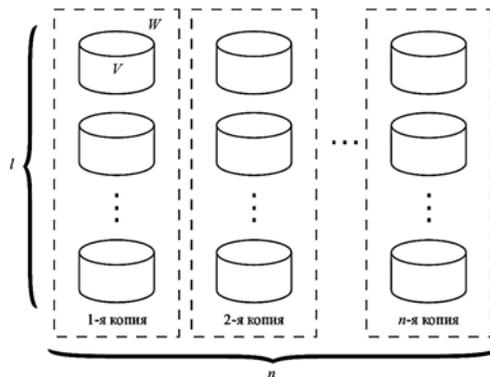


Рис. 7. Структура электронного файлового архива с несколькими копиями информации (l — число копий)

Стоимость такого решения составит

$$C_n = c \ln n + C_{\text{п.}}$$

С помощью полученной модели оценим надежность сохранения информации в электронном архиве для условий, приведенных выше. Причем с практической точки зрения имеет смысл рассмотреть только варианты реализации архивного хранилища с минимальной стоимостью (на базе BD-R) и с наилучшей надежностью (на базе BD-R XL).

Результаты расчетов для $n = 2$ приведены на рис. 8, 9 (отсутствующие части линий графиков вероятности отказа не могут быть построены из-за ограничений разрядной сетки программы octave, использовавшейся для моделирования).

Как можно видеть, для $q < 0,0001$ итоговая надежность хранения информации в архиве превосходит надежность одного оптического диска. То есть при выборе высококачественных (дорогих) дисков и определенного числа копий n может быть получено необходимое значение надежности хранения информации Q_n .

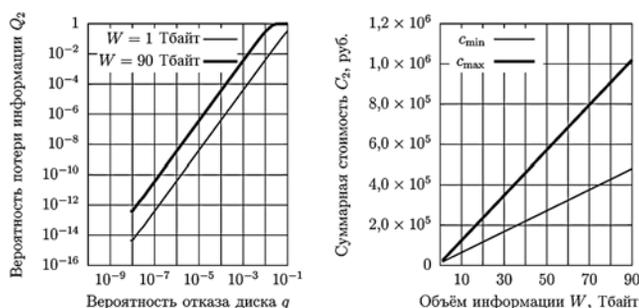


Рис. 8. Результаты расчетов для архива на базе дисков BD-R для $n = 2$

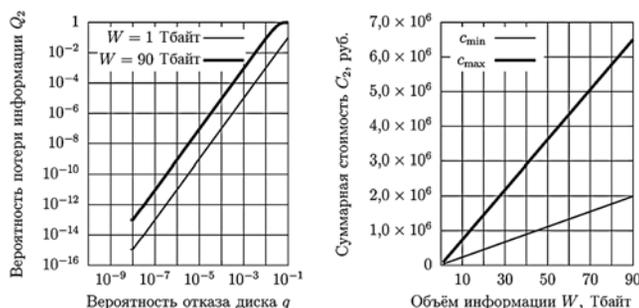


Рис. 9. Результаты расчетов для архива на базе дисков BD-R XL для $n = 2$

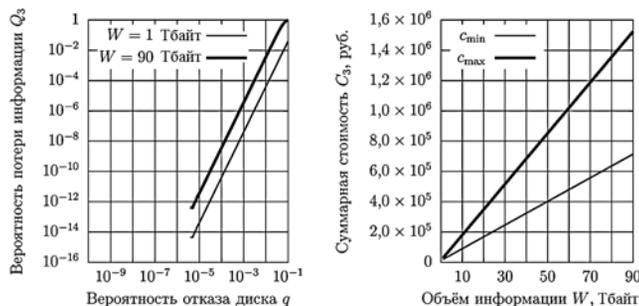


Рис. 10. Результаты расчетов для архива на базе дисков BD-R для $n = 3$

Поскольку очевидно, что для одних и тех же значений n решение с применением BD-R XL по надежности превосходит решение с применением BD-R, но существенно дороже его, представляет интерес оценка пути достижения необходимого значения Q_n при минимизации стоимости решения. Для этого выполним расчеты для решения с применением BD-R и $n = 3$ (рис. 10).

Как можно видеть из рис. 10, уже для $q < 0,01$ итоговая надежность хранения информации в архиве превосходит надежность одного оптического диска. Но главное, что для архивного хранилища на базе BD-R при $n = 3$ с уменьшением q надежность хранения информации растет значительно быстрее, чем для архивного хранилища на базе BD-R XL при $n = 2$. И при этом суммарная стоимость хранения информации для BD-R остается все еще значительно ниже (примерно в 4 раза) по сравнению с BD-R XL.

Заключение

Таким образом, на сегодняшний день при создании долговременных электронных архивных хранилищ информации в небольших организациях имеет смысл ориентироваться на диски типа BD-R в качестве носителей информации. Требуемый уровень надежности хранения может быть достигнут за счет создания необходимого числа копий n , а также использования качественных дисков от зарекомендовавших себя производителей (обычно такие диски более дороги).

Список литературы

- ГОСТ Р 54989—2012/ISO TR 18492:2005. Обеспечение долговременной сохранности электронных документов. (Вступил в силу 01.05.2013).
- Корепанов И. Как сохранить архив на десятилетия? [Электронный ресурс] // Журнал сетевых решений/ LAN. 2008, № 3. URL: <http://www.osp.ru/lan/2008/03/4899898/>. Дата обращения: 2015-11-04.
- Российские организации начинают строить электронные архивы на оптических дисках. [Электронный ресурс] // Ассоциация электронных торговых площадок. URL: <http://www.aetr.ru/market-news/item/400867>. Дата создания: 06.11.2015. Дата обращения: 13.11.2015.
- Категория: Облачные хранилища данных: Материал из Википедии — свободной энциклопедии. [Электронный ресурс]. URL: https://ru.wikipedia.org/wiki/Категория:Облачные_хранилища_данных — Дата обращения: 2015-12-25.
- Каташин В. В., Чернышов А. В. Проблема надежного хранения файлов в архивных электронных хранилищах небольших организаций (на примере музеев) // Перспективы развития информационных технологий: сборник материалов XIX Международной научно-практической конференции. Новосибирск: Изд-во ЦРНС, 2014. С. 12—23.
- Школьный музей: из века XX — в XXI. [Электронный ресурс] // Музей. 2010, № 7. С. 70—73. URL: http://turcentrf.ru/d/358156/d/001_persin-1.pdf. Дата обращения: 2015-12-26
- Новикова Е. Не последний аккорд, библиотечное дело и хор мальчиков... [Электронный ресурс]. URL: http://gazeta.zn.ua/CULTURE/ne_posledniy_akkord_bibliotechnoe_delo_i_hor_malchikov.html. Дата создания: 02.07.2004. Дата обращения: 10.12.2015.
- Устинов В. Хранение данных на CD- и DVD-дисках: на наш век хватит? [Электронный ресурс] // BROADCASTING: Телевидение и радиовещание. URL: http://www.broadcasting.ru/articles2/Oborandteh/hranenie_dannyh_na_CD_DVD_diskah. Дата обращения: 10.12.2015.
- Вентцель Е. С., Овчаров Л. А. Теория вероятностей и ее инженерные приложения. М.: Наука. Гл. ред. физ.-мат. лит. 1988. 480 с.
- Sony ODS-L10 — библиотека архива оптических дисков с 10 гнездами [Электронный ресурс] // Телерадио: Оборудование для РАДИО и ТВ. URL: <http://tele-radio.ru/product/sony-ods-l10/> Дата обращения: 13.11.2015.
- Робот-библиотекарь поселился в новом кампусе СПбГУ "Михайловская дача" [Электронный ресурс] // Петербургский

государственный университет. 2015. URL: <http://spbu.ru/smi/o-nas-pishut/24726-robot-bibliotekar-poselilsya-v-novom-kampuse-spbgu-mikhailovskaya-dacha.html>. Дата публикации: октябрь, 2015. Дата обращения: 13.11.2015.

12. **Жесткий** диск HDD SATA [Электронный ресурс] // wikimart. URL: <http://computers.wikimart.ru/components/hdd/tag/hdd-sata>. Дата обращения: 17.12.2015.

13. **Blu-ray** диски для записи [Электронный ресурс] // Store52.ru: Магазин носителей информации. 2015. URL: http://www.store52.ru/catalog/blu-ray_diski_dlya_zapisi. Дата обращения: 17.12.2015.

14. **BD-R DL** [Электронный ресурс] // Ресурс ППО. URL: <http://resurs-pro.ru/catalog/bd-r-dl>. Дата обращения: 17.12.2015.

A. V. Chernyshov, Associate Professor, e-mail: sch@mgul.ac.ru, Moscow State Forest University

To the Question of the Optical Discs Application for Long Term Digital Archive Storage of Small Organizations

In the article discuss the problem of creation of electronic archive storage of information for a small organization. Usually for electronic archival storage of such organizations does not make sense to use on the market today robotic optical disc libraries because of high cost and low efficiency of use. Cost minimization can be achieved by creating an archive on a conventional a written once optical media, accessed manually, and the required reliability of storage — by creating required number of copies.

The article is built the model in estimating the reliability of save the information specified volume in the storage, based on optical disks of different capacity (CD-R, DVD-R, BD-R, BD-R DL, BD-R XL), as well as total costs for the establishment of such a repository. It is shown that the most promising are solutions based on BD-R and BD-R XL. The target value is reached the secure storage of information while minimizing the cost is achieved when the store is created based on BD-R.

Keywords: *electronic archival storage of information, long-term storage, optical storage media, reliability of information storage, cost minimization*

References

1. **GOST R 54989—2012/ISO TR 18492:2005.** *Obespechenie dolgovremennoj sohrannosti jelektronnyh dokumentov.* [Ensuring long-term preservation of electronic records.] (Entered to force 01.05.2013).

2. **Korepanov I.** Kak sohranit' arhiv na desjatiletija? [How to save an archive on a few decades?] [Electronic resource], *Zhurnal setevyh reshenij LAN* [Network solutions magazine/LAN]. 2008, vol. 03. URL: <http://www.osp.ru/lan/2008/03/4899898/>. Access date: 2015-11-04.

3. **Rossijskie** organizacii nachinajut stroit' jelektronnye arhivy na opticheskix diskah. [Russian organizations are beginning to build electronic archives on optical discs.] [Electronic resource] (in Russian), *Associacija jelektronnyh togovyh ploshhadok* [The Association of electronic trading platforms]. URL: <http://www.aetp.ru/market-news/item/400867>. Create date: 06.11.2015. Access date: 13.11.2015.

4. **Kategorija: Oblachnye hranilishha dannyh: Material iz Vikipedii — svobodnoj jenciklopedii.** [Cloud storage: from Wikipedia, the free encyclopedia.] [Electronic resource] (in Russian). URL: https://ru.wikipedia.org/wiki/Kategorija:Oblachnye_hranilishha_dannyh. Access date: 2015-12-25.

5. **Katashin V. V., Chernyshov A. V.** Problema nadjozhnogo hranenija fajlov v arhivnyh jelektronnyh hranilishhah nebol'shih organizacij (na primere muzeev). [The problem of reliable storage of files in archival electronic storage of small organizations (on the example of museums)], *Perspektivy razvitiya informacionnyh tehnologij: sbornik materialov XIX Mezhdunarodnoj nauchno-prakticheskoj konferencii.* [Prospects of development of information technologies: proceedings of the XIX International scientific and practical conference], Novosibirsk: CRNS, 2014, pp. 12—23.

6. **Shkol'nyj muzej: iz veka XX — v XXI.** [School Museum: from the XX century — to XXI.] [Electronic resource] (in Russian), *Muzej* [Museum], 2010, vol. 7, pp. 70—73. URL: http://turcentrf.ru/d/358156/d/001_persin-1.pdf. Access date: 2015-12-26

7. **Novikova E.** *Ne poslednij akkord, bibliotechnoe delo i hor mal'chikov...* [Not the last chord, libraries and boys choir...] [Elec-

tronic resource] (in Russian). URL: http://gazeta.zn.ua/CULTURE/ne_poslednij_akkord_bibliotechnoe_delo_1_hor_malchikov.html. Create date: 02.07.2004. Access date: 10.12.2015.

8. **Ustinov V.** Hranenie dannyh na CD- i DVD-diskah: na nash vek hvatit? [Data storage on CD and DVD: on our century will suffice?] [Electronic resource] (in Russian), *BROADCASTING: Televizenie i radioveshhanie* [Television and broadcasting]. URL: http://www.broadcasting.ru/articles2/Oborandteh/hranenie_dannyh_na_CD_DVD_diskah. Access date: 10.12.2015.

9. **Ventcel' E. S., Ovcharov L. A.** *Teorija verojatnostej i ejo inzhenernye prilozhenija* [Probability theory and its engineering applications]. Moscow: Nauka. Gl. red. fiz.-mat. lit., 1988, 480 p.

10. **Sony ODS-L10** — biblioteka arhiva opticheskix diskov s 10 гнездami [Sony ODS-L10 — Archive library of optical disks with 10 sockets] [Electronic resource] (in Russian), *Teleradio: Oborudovanie dlja RADIO i TV* [Teleradio: Radio & TV equipment], URL: <http://tele-radio.ru/product/sony-ods-l10>. Access date: 13.11.2015.

11. **Robot-bibliotekar** poselilsja v novom kampuse SPbGU "Mikhailovskaya dacha" [Robot librarian settled in the new campus of SPSU "Mikhailovskaya dacha"] [Electronic resource] (in Russian), *Petersburgskij gosudarstvennyj universitet*, 2015, [St. Petersburg state University]. URL: <http://spbu.ru/smi/o-nas-pishut/24726-robot-bibliotekar-poselilsya-v-novom-kampuse-spbgu-mikhailovskaya-dacha.html>. Publ. date: October, 2015. Access date: 13.11.2015.

12. **Zhlostkij disk HDD SATA** [Hard disk with SATA interface] [Electronic resource] (in Russian). wikimart. URL: <http://computers.wikimart.ru/components/hdd/tag/hdd-sata>. Access date: 17.12.2015.

13. **Blu-ray** диски dlja zapisi [Blue-ray disks for writing] [Electronic resource] (in Russian). *Store52.ru: Magazin nositelej informacii*, 2015. [Store52.ru: Shop of information media]. URL: http://www.store52.ru/catalog/blu-ray_diski_dlya_zapisi. Access date: 17.12.2015.

14. **BD-R DL** [Electronic resource] (in Russian). 2015. URL: <http://resurs-pro.ru/catalog/bd-r-dl>. Access date: 17.12.2015.

Адрес редакции:

107076, Москва, Стромынский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор *Е. В. Конова.*

Корректор *К. В. Комиссарова.*

Сдано в набор 07.06.2016. Подписано в печать 21.07.2016. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ ИТ816. 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.