

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
«Южно-Российский государственный университет экономики и сервиса»
(ФГБОУ ВПО «ЮРГУЭС»)
Волгодонский институт сервиса (филиал)
(ВИС ФГБОУ ВПО «ЮРГУЭС»)

ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ. ТЕОРИЯ И ПРАКТИКА

Сборник научных трудов



ШАХТЫ
ФГБОУ ВПО «ЮРГУЭС»
2011

УДК 004
ББК 32.97
И741

Редакционная коллегия:

А.Н. Берёза, к.т.н., доцент (председатель редакционной коллегии);

Д.А. Безуглов, д.т.н., профессор;

Н.Е. Галушкин, д.т.н., профессор;

А.В. Коротков, д.ф.-м.н., академик МАСИ;

В.М. Курейчик, д.т.н., профессор;

В.Е. Мешков, к.т.н., профессор;

Ю.А. Валюкевич, к.т.н., профессор;

В.В. Семенов, к.т.н., доцент

И741 Информационные системы и технологии. Теория и практика :
сборник научных трудов / редколлегия : А.Н. Берёза [и др.]. – Шах-
ты : ФГБОУ ВПО «ЮРГУЭС», 2011. – 283 с.
ISBN 978-5-93834-671-0

В тематический сборник включены работы учёных, проводящих исследования в следующих областях: теоретические основы информатики, интеллектуальные информационные системы, оптоинформатика, системы автоматизации проектирования, моделирование технических систем, web-технологии, менеджмент информационных технологий, информационные технологии в образовании.

Настоящий сборник предназначен для широкого круга научных работников и специалистов, а также студентов старших курсов, магистрантов и аспирантов.

УДК 004
ББК 32.97

Тексты набраны с авторских оригиналов. Редакционная коллегия приносит извинения за возможные неточности, возникшие в процессе компьютерной вёрстки издания.

ISBN 978-5-93834-671-0

© ФГБОУ ВПО «Южно-Российский государственный университет экономики и сервиса», 2011

© ВИС (филиал) ФГБОУ ВПО «ЮРГУЭС», 2011

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	6
--------------------------	----------

Раздел 1 **ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ**

Ионин В.К. Силлогизмы для нечёткой аристотелевой логики с непрерывной и строго монотонной нормой	7
Куркина М.В., Львова М.А. Транзитивное замыкание толерантного нечёткого случайного отношения	13
Никонорова Ю.В. Применение пакета аналитических вычислений MAPLE к решению задач евклидовой геометрии.....	20
Коровин Е.Н. Применение пакета аналитических вычислений Maple при нахождении метрики и кривизны для 6-мерных нильпотентных симплектических групп Ли.....	27
Коротков А.В. Многомерные алгебры полей сравнений по mod=2	36
Коротков А.В. Многомерные Булевы алгебры.....	42

Раздел 2 **ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ**

Бегляров В.В. Систематический анализ принципов эволюционного моделирования	49
Ляшов М.В. Генетический алгоритм кодирования состояний конечного автомата	68
Козоброд А.В., Мешков В.Е., Мешкова Е.В. Обработка текстовых документов на основе законов Зипфа для формирования гибридного нейросетевого классификатора	75
Мешков В.Е., Чураков В.С. Программа исследований в области информационных технологий, искусственного интеллекта и когнитологии в рамках семимерной парадигмы А.В. Короткова.....	82

Раздел 3 **ОПТОИНФОРМАТИКА**

Бекашева З.С., Васильев В.Н., Востриков А.А., Павлов А.В. Статистические свойства гипотез, порождаемых двуслойной нейронной сетью с матрицей двунаправленных связей, реализуемой методом голографии Фурье	90
Орлов В.В. Ассоциативные связи между словами, хранящимися в памяти человека, согласно голографической модели памяти	106

Раздел 4

СИСТЕМЫ АВТОМАТИЗАЦИИ ПРОЕКТИРОВАНИЯ

Гладков Л.А., Лисовцова А.Е. Применение нечётких генетических алгоритмов с вещественным кодированием для решения задач оптимизации.....	113
Стороженко А.С., Подстрелов А.Г. Система оптимизации математических функций для вычисления на ЭВМ.....	119

Раздел 5

МОДЕЛИРОВАНИЕ ТЕХНИЧЕСКИХ СИСТЕМ

Токарев А.А., Мешков В.Е. Формализация процедуры свёртки топологически сложных участков гидравлических систем паровых котлов на основе свойств матрицы инцидентности.....	125
Семёнов В.В., Ханжонков Ю.Б., Асцатуров Ю.Г. Образцовый телевизионный анализатор промышленных аэрозолей	130
Галушкин Н.Е., Галушкина Н.Н. Структурное моделирование процесса саморазряда в щелочных аккумуляторах	138
Медведев Д.В. Алгоритм определения пропускной способности автоматизированной системы контроля и учёта потребления электроэнергии в сетях 0,4 кВ.....	149
Валюкевич Ю.А., Зеленский А.А. Программный интерполятор отрезка прямой для системы управления двухзвенным механизмом	154
Наумов И.И. Кинематическая модель манипулятора для силомоментной обработки плоских материалов.....	164
Толстунов О.Г. Моделирование пространственного манипулятора с гибким подвесом объекта перемещения в среде MATLAB.....	168
Алепко А.В. Модель троса с квазираспределёнными параметрами для пространственного манипулятора с гибким подвесом объекта перемещения	177
Козаченко П.Н., Дубовсков В.В. Программно-аппаратный комплекс цифровой обработки акустико-эмиссионных сигналов химических реакций.....	181
Безуглов Д.А., Андриющенко И.В., Швидченко С.А. Информационная технология идентификации закона распределения на базе кумулянтного метода анализа результатов измерений.....	186

Клименко Н.Б., Трясоруков А.И., Алабут А.В. Обработка исходных данных для применения методов математического моделирования в исследовании влияния особенностей коленного сустава на наличие признаков его заболеваний	194
Пархоменко Н.Г. Поляризационно-независимое оценивание углового положения источников когерентных сигналов на основе методов неквадратичной регуляризации	201
Пархоменко Н.Г. Потенциальные характеристики радиолокационных систем с посторонним «подсветом».....	213

Раздел 6 WEB-ТЕХНОЛОГИИ

Бахман Е.А., Кизим А.В. Автоматизация прикладных задач с помощью web-ориентированных систем	230
--	-----

Раздел 7 МЕНЕДЖМЕНТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Божич В.И., Савченко М.Б., Попова З.И. Концептуальная модель информационных процессов совместной деятельности	235
Савченко М.Б. Информационно-психологическая безопасность «белых воротничков».....	243
Божич В.И., Савченко М.Б. Интеллектуальные инструментальные средства менеджмента информационных процессов компьютерной педагогики	250
Попова З.И., Савченко М.Б., Божич В.И. Эмерджентный интеллект информационного менеджмента	261

РАЗДЕЛ 8 ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

Ершова Е.А. Модель иерархической нечёткой системы поддержки принятия решения в вузе	272
Сведения об авторах	280

ПРЕДИСЛОВИЕ

В настоящее время развитие информационных систем и технологий основывается на разработке новых математических и алгоритмических средств, интеллектуальных методов и моделей, совершенствовании полупроводниковых и оптических технологий.

Сборник состоит из восьми тематических разделов: «Теоретические основы информатики», «Интеллектуальные информационные системы», «Оптоинформатика», «Системы автоматизации проектирования», «Моделирование технических систем», «Web-технологии», «Менеджмент информационных технологий», «Информационные технологии в образовании».

В первый раздел включены работы, посвящённые теоретическим основам информатики. В работах известных российских учёных отражены новые научные знания о применении перспективных математических парадигм, таких как нечёткие и многомерные логики.

Во втором разделе представлены работы, предлагающие новые подходы к построению интеллектуальных информационных систем.

В третьем разделе собраны научные работы, касающиеся нового перспективного направления – оптоинформатики.

Четвёртый раздел включает работы, в которых предложены новые подходы к созданию интеллектуальных систем автоматизации проектирования.

В пятом разделе представлены работы, посвящённые моделированию технических систем различного назначения.

В шестом разделе предложена работа, решающая задачу повышения эффективности работы пользователей в сети Интернет за счёт автоматизации прикладных задач с помощью web-ориентированных систем.

В седьмом разделе представлены работы в области менеджмента информационных технологий.

В восьмом разделе представлена работа по проблемам информатизации образования и внедрению перспективных информационных технологий в управление образованием.

Настоящий сборник предназначен для широкого круга инженерных и научных работников, аспирантов, магистрантов и студентов старших курсов, научные интересы которых связаны с разработкой и использованием информационных систем и технологий, основанных на новейших принципах.

Сборник подготовлен при непосредственном участии сотрудников кафедры «Информатика» ВИС ЮРГУЭС.

Редакционная коллегия заранее благодарна за отзывы и замечания, которые следует направлять по адресу: 347375, Россия, Ростовская обл., г. Волгодонск, ул. Черникова, 6, ВИС ЮРГУЭС, кафедра «Информатика».

РАЗДЕЛ 1

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ

УДК 519.6

В.К. Ионин

СИЛЛОГИЗМЫ ДЛЯ НЕЧЁТКОЙ АРИСТОТЕЛЕВОЙ ЛОГИКИ С НЕПРЕРЫВНОЙ И СТРОГО МОНОТОННОЙ НОРМОЙ*

Введение. Основные определения

Классическая логика Аристотеля состоит из предложений вида:

$$P \alpha Q, P i Q, P e Q, P o Q,$$

где P и Q – переменные, значениями которых служат множества (классы объектов), а a, i, e и o – аристотелевы связки (кванторы), обозначающие отношения между множествами [1]. Этими отношениями служат, соответственно, включение, непустое пересечение, пустое пересечение и отрицание включения множеств P^I и Q^I , которые служат значениями переменных P и Q в интерпретации I . Таким образом, если U^I – универсум интерпретации I , то $P^I \subseteq U^I, Q^I \subseteq U^I$ и

$$(P a Q)^I = 1 \text{ (истина)} \Leftrightarrow P^I \subseteq Q^I, (P i Q)^I = 1 \Leftrightarrow P^I \cap Q^I \neq \emptyset, \\ (P e Q)^I = 1 \Leftrightarrow P^I \cap Q^I = \emptyset, (P o Q)^I = 1 \Leftrightarrow P^I - Q^I \neq \emptyset.$$

Нечёткая Аристотелева логика имеет тот же синтаксис, т.е. её предложения имеют вид $P \alpha Q$, где $\alpha \in \{a, i, e, o\}$. Интерпретация I назначает каждому имени P нечёткое подмножество P^I универсума U^I , т.е. функцию $P^I: U^I \rightarrow [0,1] = \{x | 0 \leq x \leq 1\}$. Значениями предложений в интерпретации I служат по определению:

$$(A) \begin{cases} (PaQ)^I = \sup\{P^I(u) \rightarrow Q^I(u) \mid u \in U^I\} = \sup\{\neg(P^I(u) \wedge \neg Q^I(u)) \mid u \in U^I\}, \\ (PiQ)^I = \inf\{P^I(u) \wedge Q^I(u) \mid u \in U^I\}, \\ (PeQ)^I = \sup\{\neg(P^I(u) \wedge Q^I(u)) \mid u \in U^I\}, \\ (PoQ)^I = \inf\{P^I(u) \wedge \neg Q^I(u) \mid u \in U^I\}. \end{cases}$$

* Работа выполнена при финансовой поддержке РФФИ (проекты 08-01-00587, 09-01-00465).

Здесь $\neg$ и $\wedge$ – нечёткие логические связки, которые обычно определяются следующим образом:

$$\neg x = 1 - x \text{ и } x \wedge y = T(x, y),$$

где T – треугольная норма [2]. В этом случае соотношения (А) переходят в соотношения:

$$(B) \begin{cases} (PaQ)^I = \sup\{1 - T(P^I(u), 1 - Q^I(u)) \mid u \in U^I\}, \\ (PiQ)^I = \inf\{T(P^I(u), Q^I(u)) \mid u \in U^I\}, \\ (PeQ)^I = \sup\{1 - T(P^I(u), Q^I(u)) \mid u \in U^I\}, \\ (PoQ)^I = \inf\{T(P^I(u), 1 - Q^I(u)) \mid u \in U^I\}. \end{cases}$$

В работе [3] было дано определение нечётких силлогизмов как правил вывода в нечёткой Аристотелевой логике.

Выражения вида $r \leq P \alpha Q \leq s$ будем называть (нечёткими) *фактами*. Здесь $\alpha \in \{a, i, e, o, a^{-1}, o^{-1}\}$, r и s – числа из интервала $[0, 1]$, причём $r \leq s$, а α^{-1} обозначает отношение, обратное к отношению α . Значением факта $r \leq P \alpha Q \leq s$ в интерпретации I служит 1 (истина) или 0 (ложь) и

$$\begin{aligned} (r \leq P \alpha Q \leq s)^I &= 1 \Leftrightarrow_{\text{def}} r \leq (P \alpha Q)^I \leq s, \\ (r \leq P \alpha^{-1} Q \leq s)^I &= 1 \Leftrightarrow_{\text{def}} r \leq (Q \alpha P)^I \leq s. \end{aligned}$$

Пусть S – множество фактов и ϕ – какой-либо факт. *Логическое следствие* $S \models \phi$ определяется стандартно, т.е. как отсутствие интерпретации, при которой все факты из S истинны, но факт ϕ ложен.

Рассмотрим правила вывода для фактов, имеющие вид:

$$R(g, h): a \leq P \alpha Q \leq c, b \leq P \beta Q \leq d \mid - g \leq P \gamma Q \leq h,$$

где a, b, c, d, g, h рассматриваются как параметры (переменные), принимающие значения в интервале $[0, 1]$.

Правило $R(g, h)$ состоятельно, если имеет место логическое следствие:

$$a \leq P \alpha Q \leq c, b \leq P \beta Q \leq d \mid = g \leq P \gamma Q \leq h.$$

Ясно, что это логическое следствие тривиально имеет место, если посылки правила $R(g, h)$ противоречат друг другу, т.е. не существует интерпретации I такой, что выполнены оба неравенства:

$$a \leq (P \alpha Q)^I \leq c \text{ и } b \leq (P \beta Q)^I \leq d. \quad (1).$$

Естественно исключить из рассмотрения такие правила вывода.

Существует условие, накладываемое на числа a, b, c и d , когда пара неравенств (1) непротиворечива. Это условие назовём *условием непротиворечивости*.

Ясно, что для любых связок $\alpha, \beta, \gamma \in \{a, i, e, o, a^{-1}, o^{-1}\}$ и любых чисел $a, b, c, d \in [0,1]$ таких, что $a \leq c$ и $b \leq d$, найдутся числа g и h такие, что правило $R(g,h)$ состоятельно (например, можно взять $g = 0$ и $h = 1$). Ясно также: если правило $R(g,h)$ состоятельно, то состоятельно тоже и правило $R(g',h')$, где $g' \leq g$ и $h' \geq h$.

Следовательно, существуют оптимальные границы для g и h , при которых состоятельно правило вывода $R(g,h)$.

Положим:

$$g_{\alpha\beta\gamma}(a,b,c,d) = \sup\{g \mid \text{правило } R(g,1) \text{ состоятельно}\},$$

$$h_{\alpha\beta\gamma}(a,b,c,d) = \inf\{h \mid \text{правило } R(0,h) \text{ состоятельно}\}.$$

Если a, b, c, d считать *параметрами*, т.е. переменными со значениями в интервале $[0,1]$, причём такими, что $a \leq c$ и $b \leq d$, то $g_{\alpha\beta\gamma}(a,b,c,d)$ и $h_{\alpha\beta\gamma}(a,b,c,d)$ будут функциями от этих параметров. Правило вывода

$$a \leq P \alpha Q \leq c, \quad b \leq Q \beta R \leq d \mid -g_{\alpha\beta\gamma}(a,b,c,d) \leq P \gamma R \leq h_{\alpha\beta\gamma}(a,b,c,d) \quad (2)$$

мы называем *нечётким силлогизмом с паттерном $\alpha \beta \gamma$* . Построить нечёткий силлогизм с паттерном $\alpha \beta \gamma$ – найти выражения для функций $g_{\alpha\beta\gamma}(a,b,c,d)$ и $h_{\alpha\beta\gamma}(a,b,c,d)$, а также указать условие непротиворечивости для правой части (2).

В настоящей работе мы рассматриваем *случай пропозициональной Аристотелевой логики*, когда допускаются только интерпретации с одноэлементными универсумами. Точнее, пусть $U^I = \{\omega\}$ для всех допустимых интерпретаций. Тогда соотношения (B) превращаются в соотношения:

$$(P \ a \ Q)^I = 1 - T(P^I(\omega), 1 - Q^I(\omega)), \quad (P \ i \ Q)^I = T(P^I(\omega), Q^I(\omega)),$$

$$(P \ e \ Q)^I = 1 - T(P^I(\omega), Q^I(\omega)), \quad (P \ o \ Q)^I = T(P^I(\omega), 1 - Q^I(\omega)).$$

Значение $P^I(\omega) \in [0,1]$ есть степень принадлежности элемента ω к нечёткому подмножеству P^I универсума U^I . Мы можем трактовать значение $P^I(\omega)$ как значение нечёткой пропозициональной переменной, которую обозначим p .

Таким образом, мы приходим к *нечёткой пропозициональной Аристотелевской логике*, предложения которой имеют вид: $paq, p i q, req, roq$, для которых имеем:

$$(p \ a \ q)^I = 1 - T(p^I, 1 - q^I), \quad (p \ i \ q)^I = T(p^I, q^I),$$

$$(p \ e \ q)^I = 1 - T(p^I, q^I), \quad (p \ o \ q)^I = T(p^I, 1 - q^I).$$

В нечёткой пропозициональной Аристотелевой логике силлогизмы принимают вид:

$$a \leq p \alpha q \leq c, b \leq q \beta r \leq d \vdash g_{\alpha\beta\gamma}(a,b,c,d) \leq p \gamma r \leq h_{\alpha\beta\gamma}(a,b,c,d).$$

Основной результат настоящей статьи состоит в теореме, которая даёт общий метод для построения нечётких пропозициональных силлогизмов для Аристотелевой логики с треугольной нормой, которая является непрерывной и строго монотонной.

1. Задача построения нечётких пропозициональных силлогизмов как задача оптимизации

Напомним, что построить нечёткий пропозициональный силлогизм с паттерном $\alpha\beta\gamma$ – значит найти выражения $g_{\alpha\beta\gamma}(a,b,c,d)$ и $h_{\alpha\beta\gamma}(a,b,c,d)$ от параметров a, b, c, d , имеющие значения супремума и инфимума чисел g и h (соответственно) таких, что справедливы логические следствия:

$$\begin{aligned} a \leq p \alpha q \leq c, b \leq q \beta r \leq d & \models g \leq p \gamma r, \\ a \leq p \alpha q \leq c, b \leq p \beta q \leq d & \models p \gamma r \leq h. \end{aligned}$$

Это означает, что при любой интерпретации I справедливы утверждения:

$$a \leq (p \alpha q)^I \leq c, b \leq (q \beta r)^I \leq d \Rightarrow g \leq (p \gamma r)^I \leq h.$$

Обозначая $p^I = x, q^I = t$ и $r^I = y$, получаем:

$$a \leq x \alpha t \leq c, b \leq t \beta y \leq d \Rightarrow g \leq x \gamma y \leq h, \quad (11)$$

где α, β и γ обозначают на этот раз соответствующие функции $[0,1]^2 \rightarrow [0,1]$, т.е.

$$\begin{aligned} x \mathbf{a} y &= 1 - T(x, 1 - y), \quad x \mathbf{i} y = T(x, y), \quad x \mathbf{e} y = 1 - T(x, y), \\ x \mathbf{o} y &= T(x, 1 - y), \quad x \mathbf{a}^* y = 1 - T(1 - x, y), \quad x \mathbf{o}^* y = T(1 - x, y). \end{aligned}$$

Введём следующие условия для пар $(x, y) \in [0,1]^2$:

$$\begin{aligned} \Delta_1 &= \Delta_1(a,b,c,d) = \Delta_1(a,b,c,d; x,y): \exists t [a \leq T(x,t) \leq c \wedge b \leq T(1-t,y) \leq d], \\ \Delta_2 &= \Delta_2(a,b,c,d) = \Delta_2(a,b,c,d; x,y): \exists t [a \leq T(x,t) \leq c \wedge b \leq T(t,y) \leq d]. \end{aligned}$$

Обозначим также:

$$\begin{aligned}
I_1 &= I_1(a,b,c,d) = \inf\{T(x,y) \mid \Delta_1(a,b,c,d)\}, \\
I_2 &= I_2(a,b,c,d) = \inf\{T(x,1-y) \mid \Delta_1(a,b,c,d)\}, \\
I_3 &= I_2(a,b,c,d) = \inf\{T(1-x,1-y) \mid \Delta_1(a,b,c,d)\}, \\
I_4 &= I_3(a,b,c,d) = \inf\{T(x,y) \mid \Delta_2(a,b,c,d)\}, \\
I_5 &= I_5(a,b,c,d) = \inf\{T(x,1-y) \mid \Delta_2(a,b,c,d)\}, \\
I_6 &= I_5(a,b,c,d) = \inf\{T(1-x,1-y) \mid \Delta_2(a,b,c,d)\}, \\
S_1 &= S_1(a,b,c,d) = \sup\{T(x,y) \mid \Delta_1(a,b,c,d)\}, \\
S_2 &= S_2(a,b,c,d) = \sup\{T(x,1-y) \mid \Delta_1(a,b,c,d)\}, \\
S_3 &= S_3(a,b,c,d) = \sup\{T(1-x,1-y) \mid \Delta_1(a,b,c,d)\}, \\
S_4 &= S_3(a,b,c,d) = \sup\{T(x,y) \mid \Delta_2(a,b,c,d)\}, \\
S_5 &= S_5(a,b,c,d) = \sup\{T(x,1-y) \mid \Delta_2(a,b,c,d)\}, \\
S_6 &= S_6(a,b,c,d) = \sup\{T(1-x,1-y) \mid \Delta_2(a,b,c,d)\}.
\end{aligned}$$

Теорема 1. Оценки $g_{\alpha\beta\gamma}$ и $h_{\alpha\beta\gamma}$ для нечётких силлогизмов выражаются через I_j, S_j ($j=1,2,\dots,6$). Точнее, $g_{\alpha\beta\gamma}(a,b,c,d)$ и $h_{\alpha\beta\gamma}(a,b,c,d)$ совпадают с выражениями вида $I_j(a',b',c',d')$, $S_j(a',b',c',d')$, $1-I_j(a',b',c',d')$, $1-S_j(a',b',c',d')$, где $a', b', c', d' \in \{a, b, c, d, 1-a, 1-b, 1-c, 1-d\}$.

2. Выражения для I_1, I_2, S_1 и S_2

В следующей теореме предполагается, что T – произвольная непрерывная и строго монотонная треугольная норма, а $E(u,v)$, где $u, v \in [0,1]$, обозначает множество $\{(x,y) \mid \exists t \in [u,v]. T(x,t) = u \wedge T(t,y) = v\}$. Также через $T(E(u,v))$ обозначено множество $\{T(x,y) \mid (x,y) \in E(u,v)\}$.

Теорема 2. Величины I_1, I_2, S_1 и S_2 выражаются так:

$$\begin{aligned}
I_1 &= 1, \text{ если } a+b = 1 \text{ и } b>0 \text{ или } a+b>1; \\
I_1 &= 0, \text{ если } ab = 0; \\
I_1 &= \min T(E(a,b)), \text{ если } a+b < 1 \text{ и } ab > 0; \\
I_2 &= 1, \text{ если } a+b > 1; \\
I_2 &= 0, \text{ если } a+b \leq 1; \\
S_1 &= 0, \text{ если } a+b > 1; \\
S_1 &= 1, \text{ если } a+b \leq 1 \text{ и } c+d \geq 1; \\
S_1 &= d, \text{ если } c = 0; \\
S_1 &= c, \text{ если } d = 0; \\
S_1 &= \max T(E(c,d)), \text{ если } c+d < 1 \text{ и } cd > 0; \\
S_2 &= 0, \text{ если } a+b > 1; \\
S_2 &= 1-b, \text{ если } a = 0; \\
S_2 &= 1, \text{ если } b = 0; \\
S_2 &= 1-b_0, \text{ если } a+b < 1, ab > 0, \text{ где } b_0 \text{ вычисляется из условия } T(1-a, b_0).
\end{aligned}$$

Другие величины $I_3, I_4, I_5, I_6, S_3, S_4, S_5$ и S_6 имеют аналогичные выражения, но в некоторые из этих выражений входит уже множество $\{(x,y) \mid \exists t \in [u,v]. T(x,t) = u \wedge T(t,y) = v\}$.

Мы видим из теоремы 1, что от нормы T зависят только значения I_1 при $a + b < 1, ab > 0$ и значения S_1 при $c + d < 1$ и $cd > 0$. Таким образом, при построении нечётких силлогизмов для логики, основанной на заданной норме T , нужно находить минимум и максимум чисел из множеств вида $T(E(u,v))$.

Возьмём мультипликативную норму $T(x,y) = xy$. Тогда, например, нечётким силлогизмом с паттерном **aii** является:

$$g_{aii} = \left. \begin{array}{l} cd, \text{ если } a = 1 \text{ или } b = 1; \\ 0, \text{ если } c = 0 \text{ или } d = 0; \\ 1, \text{ если } b \leq c, a \leq d; \\ \frac{c}{d}, \text{ если } 0 < c < b < 1; \\ \frac{d}{a}, \text{ если } 0 < d < a < 1. \end{array} \right\}$$

$$h_{aii} = \left. \begin{array}{l} 1, \text{ если } a+b \geq 1, a > 0, b > 0; \\ 4ab, \text{ если } a \leq b \leq \frac{1}{2}. \end{array} \right\}$$

Список литературы

1. Stanford Encyclopedia of Philosophy. Aristotle's Logic. – URL: <http://plato.stanford.edu/aristotle-logic>.
2. Новак В., Перфильева И., Мачкорж И. Математические принципы нечёткой логики. – М.: Физматлит, 2006.
3. Ионин В.К., Плесневич Г.С. Нечёткие пропозициональные силлогизмы // Интегрированные модели и мягкие вычисления в искусственном интеллекте: сб. науч. трудов V-й Междунар. науч.-практич. конф. – М.: Физматлит. – Т. 1. – С. 52–62.

УДК 519.7

М.В. Куркина, М.А. Львова

ТРАНЗИТИВНОЕ ЗАМЫКАНИЕ ТОЛЕРАНТНОГО НЕЧЁТКОГО СЛУЧАЙНОГО ОТНОШЕНИЯ

При социологических или экономических исследованиях часто возникает задача классификации объектов. В настоящее время в подобных задачах успешно используются методы и понятия нечёткой теории множеств [1, 2, 5]. Одно из ключевых понятий теории нечётких множеств – нечёткое отношение толерантности. В данной работе обсуждаются проблемы, связанные с транзитивным замыканием нечёткого отношения толерантности и возникающих статистических закономерностей.

Работа выполнена при финансовой поддержке РФФИ (№ 08-01-98001), Совета по грантам Президента РФ для поддержки молодых учёных и ведущих научных школ РФ (№ НШ-5682.2008.1), а также при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 гг. (гос. контракт № 02.740.11.0457).

Пусть X – конечное множество объектов.

Определение. Нечётким толерантным отношением на X называется функция $\theta: X \times X \rightarrow [0, 1]$, обладающая свойствами:

$$\begin{aligned}\theta(x, y) &= \theta(y, x), \\ x = y &\Rightarrow \theta(x, y) = 1.\end{aligned}$$

Величина $\theta(x, y)$ интерпретируется как мера сходства объектов $x, y \in X$, величина $r(x, y) = 1 - \theta(x, y)$, соответственно, как мера несходства объектов.

Определение. Нечёткое толерантное отношение на X называется нечётким отношением эквивалентности, если дополнительно выполняется аксиома транзитивности:

$$\theta(x, y) \geq \min\{\theta(x, y), \theta(y, z)\} \quad \forall x, y, z \in X.$$

Понятие нечёткого отношения эквивалентности тесно связано с хорошо известным понятием ультраметрики. Неотрицательная функция $r(x, y) = 1 - \theta(x, y)$ является ультраметрикой, если выполняется:

$$\begin{aligned}r(x, y) &= r(y, x), \\ x = y &\Rightarrow r(x, y) = 0, \\ r(x, y) &\leq \max\{r(x, y), r(y, z)\} \quad \forall x, y, z \in X.\end{aligned}$$

Ультраметрические пространства имеют важные приложения в задачах классификации и обладают замечательными свойствами [3, 4].

Определение. Транзитивным замыканием нечёткого отношения толерантности θ на множестве X называется наименьшее нечёткое отношение эквивалентности $\bar{\theta}$ такое, что $\bar{\theta} \supseteq \theta$, т.е.

$$\bar{\theta}(x, y) \geq \theta(x, y) \quad \forall x, y \in X,$$

где $\bar{\theta}$ – нечёткое отношение эквивалентности,

если θ_1 обладает свойствами 1 и 2, то $\theta_1 \geq \bar{\theta}$.

В случае конечного множества X (дискретного компактного) имеется стандартная процедура построения транзитивного замыкания, именно:

$$\bar{\theta} = \theta \cup \theta^2 \cup \theta^3 \cup \dots,$$

где отношение $\theta^k = \theta \circ \theta \circ \dots \circ \theta$ получено из отношения θ k -кратной $\max$ – $\min$ композицией. Композиция $\max$ – $\min$ определяется формулой:

$$\theta_1 \circ \theta_2(x, y) = \max_z \min\{\theta_1(x, z), \theta_2(z, y)\}.$$

Справедлива теорема.

Теорема 1. Если θ – нечёткое толерантное отношение на конечном множестве X , содержащем N элементов, то:

$$1. \quad \bar{\theta} = \theta^{N-1}.$$

2. Функция $\bar{\theta}: X \times X \rightarrow [0, 1]$ принимает не более $N-1$ различных значений, отличных от 1 значений. Обозначим это число через $q(\bar{\theta}) \leq N-1$.

Доказательство. Первое утверждение хорошо известно. Докажем второе. Заметим, что для любых $x, y, z \in X$ хотя бы два значения функции принадлежности θ нечёткой эквивалентности θ из трёх значений – $\theta(x, y)$, $\theta(y, z)$, $\theta(z, x)$ – совпадают. Пусть, для определённости

$$\theta(x, y) \geq \theta(y, z) \geq \theta(z, x).$$

Так как θ является нечётким отношением эквивалентности, то оно удовлетворяет свойствам симметричности и транзитивности, поэтому имеем:

$$\begin{aligned} \theta(z, x) &= \theta(x, z); \\ \theta(x, z) &\geq \min(\theta(x, y), \theta(y, z)). \end{aligned}$$

Согласно допущению, $\theta(y, z) \leq \theta(x, y)$. Поэтому $\min(\theta(x, y), \theta(y, z)) = \theta(y, z)$. Из условий 1, 2 получаем $\theta(z, x) \geq \theta(y, z)$. С другой стороны, имеем $\theta(z, x) \leq \theta(y, z)$. Значит, $\theta(z, x) = \theta(y, z)$.

Следовательно, при $N = 3$ утверждение верно. Пусть теорема верна для всех $N \leq K$. Покажем, что тогда она верна и для $N = K + 1$.

Выберем произвольный элемент $x_0 \in X$ и рассмотрим сужение R_0 нечёткой эквивалентности R с областью определения $(X \setminus \{x_0\}) \cdot (X \setminus \{x_0\})$. Число различных, отличных от 1, значений функции принадлежности μ_{R_0} не больше $K - 1$ (по предположению индукции). Пусть существует $x_1 \in (X \setminus \{x_0\})$ такой, что $\theta(x_0, x_1)$ не равно никакому из этих $K - 1$ значений. Тогда, по лемме, для любого $y \in (X \setminus \{x_0\})$, или $\theta(x_0, y) = \theta(x_0, x_1)$, или $\theta(x_0, y) = \theta(y, x_1)$. Следовательно, число различных, отличных от 1, значений функции принадлежности θ не больше K .

Замечание. Так как понятия нечёткого отношения эквивалентности и ультраметрики двойственные, то для каждого утверждения о нечётком отношении эквивалентности имеется соответствующее утверждение об ультраметрике, и обратно.

Матрица нечёткого отношения эквивалентности (ультраметрики) может быть приведена к специальному виду:

Лемма. Пусть $r : X \cdot X \rightarrow R$ – ультраметрика на конечном множестве X , состоящем из N элементов, тогда матрицу значений $R = \|r_{i,j}\|_{i,j=1,\dots,N}$ ультраметрики путём перенумерации элементов в X можно привести к виду:

$$R = \begin{bmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \end{bmatrix},$$

где $R_{1,1}$ и $R_{2,2}$ – квадратные подматрицы, а подматрица $R_{1,2} = R_{2,1}^T$ состоит из одинаковых чисел $\alpha_0 = \max_{i,j} r_{i,j}$.

Доказательство леммы. Введём на множестве X отношение $x \sqsubseteq_{\alpha_0} y \equiv r(x, y) < \alpha_0$, легко проверить, что это есть отношение эквивалентности (рефлексивное, симметричное, транзитивное). Множество X разбивается на классы эквивалентности:

$$X = X_1 \cup \dots \cup X_s.$$

Перенумеруем элементы множества X так, чтобы первые k -элементов составляли множество X_1 , а последующие $(N - k)$ элементов составляли множество $X_2 \cup \dots \cup X_s$. В результате матрица R примет иско-мый вид.

Замечание. Применяя последовательно эту лемму к подмножествам X_1 и $X_2 \cup \dots \cup X_s$, матрицу ультраметрики путём перенумерации элемен-тов множества X можно привести к стандартной форме. Пример матрицы 7×7 в стандартной форме приведён на рисунке 1.

0	0.7	0.8	0.9			
0.7	0					
0.8						
0.9			0	0.7	0.7	
			0.7	0		
			0.7		0	0.5
					0.5	0

Рис. 1. Стандартная форма матрицы ультраметрики

При практических социологических или экономических обследо-ваниях матрица толерантного отношения определяется эмпирически и, сле-довательно, неточно. В связи с этим возникают интересные проблемы.

Проблема 1. Исследование статистических закономерностей в рас-пределении значений коэффициентов матрицы $\bar{\theta}$ в зависимости от рас-пределения значений исходной матрицы θ . Например, если коэффициенты θ (вне главной диагонали) равномерно и независимо распределены в от-резке $[0,1]$, какие выводы можно сделать об $\bar{\theta}$?

Рассмотрим для примера случай, когда множество X состоит из 4 объектов.

Теорема 2. Пусть толерантное отношение θ имеет вероятностное распределение Бернулли, т.е. $\theta(x_i, x_i) = 1$ и при $i \neq j$, тогда

$$\theta(x_i, x_j) = \begin{cases} 1 & \text{с вероятностью } p \\ 0 & \text{с вероятностью } q = 1-p \end{cases}$$

Случайные величины $\theta(x_i, x_j)$ независимы. Тогда вероятности P_k того, что транзитивное замыкание $\bar{\theta}$ имеет k классов эквивалентности, равны:

$$\begin{aligned} P_1 &= p^3(-6p^3 + 24p^2 - 33p + 16), \\ P_2 &= (1-p)^3 p^2(15-11p), \\ P_3 &= 6(1-p)^5 p, \\ P_4 &= (1-p)^6. \end{aligned}$$

Доказательство. Другими словами, рассмотрим случайный граф в смысле определений Пола Эрдоса и Альфреда Реньи (1959) [6, 7].

1. Вероятность того, что $\bar{\theta}$ состоит из одного класса эквивалентности. Далее будем рассматривать отношения толерантности как графы (в нашем случае) с четырьмя вершинами. Чтобы транзитивное замыкание имело один класс эквивалентности, необходимо и достаточно, чтобы граф отношения толерантности был связным. Связный граф с минимальным количеством рёбер – это дерево. На t вершинах можно построить t^{t-2} различных деревьев. На четырёх вершинах – 4^2 , при этом такие деревья будут иметь три ребра. Итого получается вероятность появления такого отношения $4^2 \cdot p^3 \cdot (1-p)^3$. Графы с четырьмя или более рёбрами все будут связными. Вероятность появления соответствующих отношений толерантности:

$$C_6^4 \cdot p^4 \cdot (1-p)^2 + C_6^5 \cdot p^5 \cdot (1-p) + C_6^6 \cdot p^6,$$

соответственно, вероятность того, что $\bar{\theta}$ состоит из одного класса эквивалентности:

$$P_1 = 4^2 \cdot p^3 \cdot (1-p)^3 + C_6^4 \cdot p^4 \cdot (1-p)^2 + C_6^5 \cdot p^5 \cdot (1-p) + C_6^6 \cdot p^6.$$

2. Вероятность того, что $\bar{\theta}$ распадётся на два класса. Графы с тремя рёбрами и двумя связными компонентами. Их число $C_6^3 - 4^2$. Графы с двумя рёбрами и двумя связными компонентами. $\frac{1}{2} C_4^2$ – число графов, где одна компонента состоит из двух вершин, вторая – тоже из двух, $C_4^1 \cdot 3^1$ – число графов, где одна компонента состоит из одной вершины, вторая – из трёх.

Соответственно, вероятность того, что $\bar{\theta}$ распадется на два класса, равна:

$$P_2 = (C_6^3 - 4^2) \cdot p^3 \cdot (1-p)^3 + \frac{1}{2} C_4^2 \cdot p^2 \cdot (1-p)^4 + 3C_4^1 p^2 (1-p)^4.$$

3. Вероятность того, что $\bar{\theta}$ распадется на три класса:

$$P_3 = C_6^1 p (1-p)^5.$$

4. Вероятность того, что $\bar{\theta}$ распадется на четыре класса. Это возможно, только если все недиагональные элементы – нули. Соответственно, эта вероятность равна:

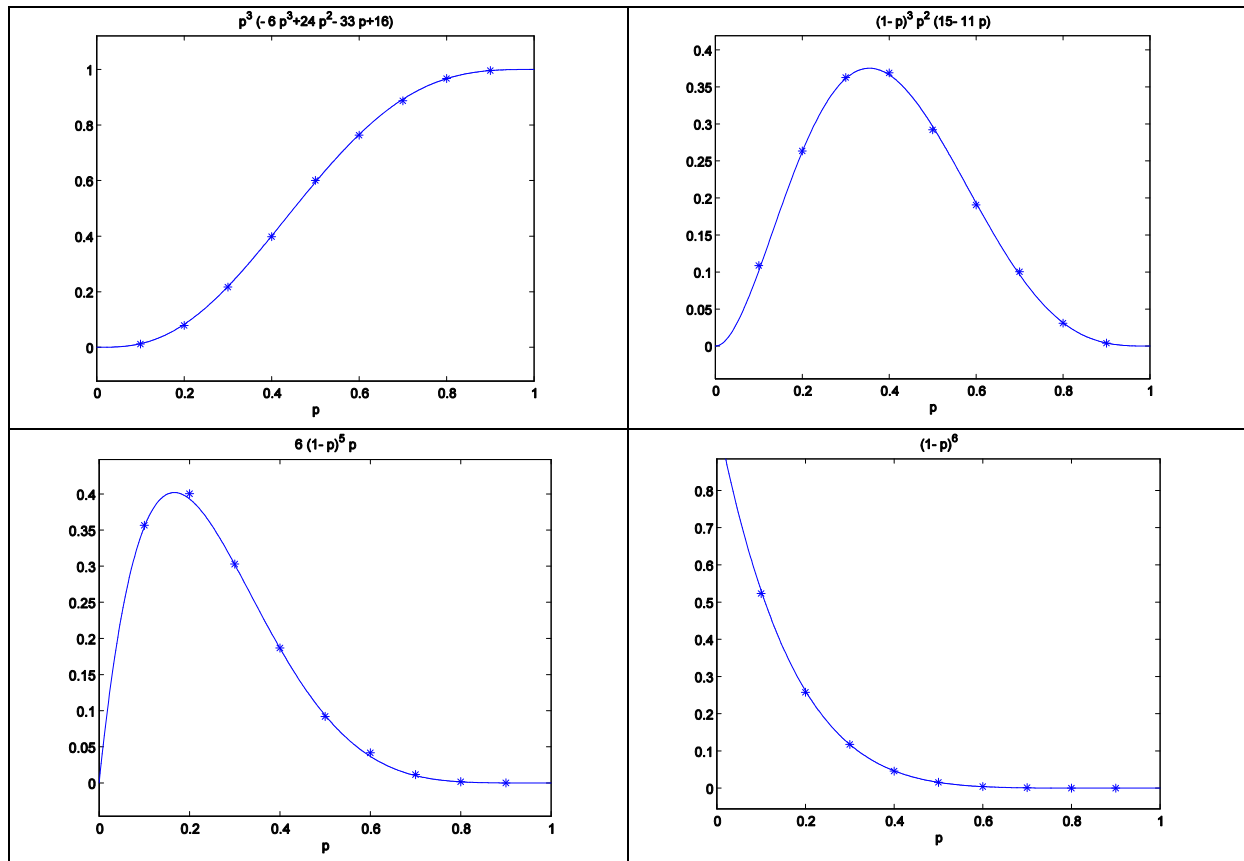
$$P_4 = (1-p)^{\frac{(4-1)4}{2}} = (1-p)^6.$$

Экспериментальные результаты

Были проведены численные эксперименты, в системе MatLab была создана m -функция **exper1(n,p,N)** (n – размерность матрицы, p – вероятность появления 1, N – число матриц), которая случайным образом генерирует N матриц отношения толерантности, ищет их транзитивное замыкание и вычисляет относительные частоты того, что транзитивное замыкание отношения распадется на определенное число классов. На выходе получается строка, в которой указаны относительные частоты, начиная с одного класса, заканчивая n классами. Для $N = 10000$ получена следующая таблица:

p	P_1	P_2	P_3	P_4
0.1	0.0117	0.1087	0.3565	0.5231
0.2	0.0787	0.2635	0.4002	0.2576
0.3	0.2168	0.3627	0.3028	0.1177
0.4	0.3984	0.3689	0.1868	0.0459
0.5	0.6004	0.2920	0.0919	0.0157
0.6	0.7635	0.1905	0.0416	0.0044
0.7	0.8870	0.1003	0.0115	0.0012
0.8	0.9673	0.0309	0.0016	0.0002
0.9	0.9960	0.0040	0.0	0.0

На рисунке 2 представлены графики функций $P_i(p)$, согласно теореме 3, и результаты эксперимента.

Рис. 2. Вероятности P_1 , P_2 , P_3 , P_4

Данные эксперимента хорошо согласуются с теоретически полученными формулами. Это даёт основание для использования m -функции **exper1(n,p,N)** в дальнейших исследованиях при $n > 4$. Данный факт важен, так как, вероятно, в общем случае получить формулы для $P_k(p)$ будет затруднительно и следует воспользоваться численным моделированием при исследовании транзитивного замыкания толерантного нечёткого случайного отношения.

Список литературы

1. Zadeh L.A. Fuzzy sets // Information and Control. – 1965. – Vol. 8. – Pp. 338–353.
2. Yu LI, Jonathan LI, Haibin DONG, Xiangqian GAO. A fuzzy relation based algorithm for segmenting color aerial images of urban environment.
3. Hazewinkel M. Classification in Mathematics, Discrete Metric Spaces, and Approximation by Trees // Report AM-R9505, CWI Amsterdam. – The Netherlands.

4. Berestovskii V.N. Ultrametric spaces. Proceedings on analysis and geometry // International conference in honor of the 70th birthday of Professor Yu.G. Reshetnyak, Novosibirsk, Russia, August 30-September 3, 1999. – Novosibirsk: Izdatelstvo Instituta Matematiki Im. S.L. Soboleva SO RAN. – 47–72 (2000).
5. Batyrshin I.Z., Rudas T., Klimova A. On general scheme of invariant clustering procedures based on fuzzy similarity relation // International Conference on Fuzzy Sets and Soft Computing in Economics and Finance FSSCEF 2004 Proceedings Volume I, Saint-Petersburg, Russia, June 17–20, 2004. – P. 122–130.
6. Колчин В.Ф. Случайные графы. – 2-е изд. – М.: ФИЗМАТЛИТ, 2004. – 256 с.
7. David J. Marchette. Random graphs for statistical pattern recognition. – Wiley-interscience, 2004.

УДК 514:004.9

Ю.В. Никонорова

ПРИМЕНЕНИЕ ПАКЕТА АНАЛИТИЧЕСКИХ ВЫЧИСЛЕНИЙ MAPLE К РЕШЕНИЮ ЗАДАЧ ЕВКЛИДОВОЙ ГЕОМЕТРИИ

В последние годы в научной среде и в математическом образовании приобрели большую популярность такие программные средства, как MathCad, MatLab, Mathematica и Maple. Указанные программные продукты позволяют освободиться от громоздких вычислений и сосредоточиться на аналитической обработке математических данных.

В настоящей статье рассмотрены примеры задач евклидовой геометрии, при решении которых удалось успешно воспользоваться возможностями пакета аналитических вычислений Maple [2–4, 9]. Все эти задачи имеют достаточно простую постановку, хотя некоторые из них, например задача Т. Поповичи (Т. Popoviciu) и задача Дж.В. Фике (J.W. Fickett), и не находили своего решения в течение долгого времени. Так, задача Т. Поповичи была поставлена в 1943 г., задача Фике была сформулирована в 80-е гг. XX в. Объединяет все рассматриваемые в настоящей статье задачи также и тот факт, что при их решении вычисления получаются настолько громоздкими, что оперировать с ними вручную практически невозможно.

Исследование каждой из представленных задач проводилось по следующему плану. Первоначально проводились многочисленные проверки соответствующей гипотезы с помощью системы Maple. То есть выбирались конкретные значения параметров, при подстановке которых и решалась задача. Затем создавалась удобная для вычислительной работы модель соответствующей задачи, при этом применялись декартовы модели евклидовой плоскости или пространства. С использованием традиционных

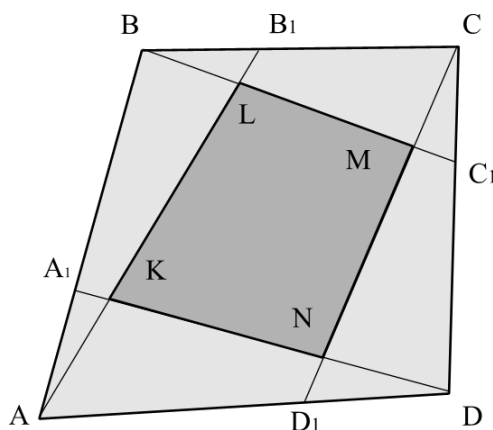
алгебраических и геометрических методов производилось упрощение задачи. Затем строился ряд алгебраических задач, соответствующих каждому шагу геометрической задачи. Для решения алгебраических задач в системе Maple приходилось либо пользоваться уже существующими процедурами поиска решения (например, стандартными процедурами для разложения на множители, упрощения выражений, построения графиков), либо создавать собственные (например более удобные процедуры решения для решения систем линейных уравнений и для вычисления определителей). Последним этапом в решении каждой из рассматриваемых задач была проверка выкладок, полученных с помощью Maple. Надо отметить, что некоторые из выкладок было трудно предугадать, и можно сказать, что система Maple «подказала» дальнейший путь решения.

1. Задача Поповичи (Т. Popoviciu)

Рассмотрим на евклидовой плоскости выпуклый четырёхугольник $ABCD$. Пусть A_1, B_1, C_1, D_1 – такие точки отрезков AB, BC, CD и DA , соответственно, что

$$\frac{|AA_1|}{|A_1B|} = \frac{|BB_1|}{|B_1C|} = \frac{|CC_1|}{|C_1D|} = \frac{|DD_1|}{|D_1A|} = k$$

для некоторого $k > 0$.



Прямые AB_1, BC_1, CD_1, DA_1 образуют четырёхугольник $KLMN$ (K, L, M, N – точки пересечения первой и четвёртой, первой и второй, второй и третьей, третьей и четвёртой прямых соответственно), лежащий внутри $ABCD$. Обозначим через S и s площади четырёхугольников $ABCD$ и $KLMN$ соответственно. Задача Т. Поповичи состояла в доказательстве неравенства при $k = 1$. При помощи компьютерных расчётов, произведённых в Maple, было получено даже более общее утверждение, справедливое для всех $k > 0$:

$$\frac{1}{6}S \leq s \leq \frac{1}{5}S.$$

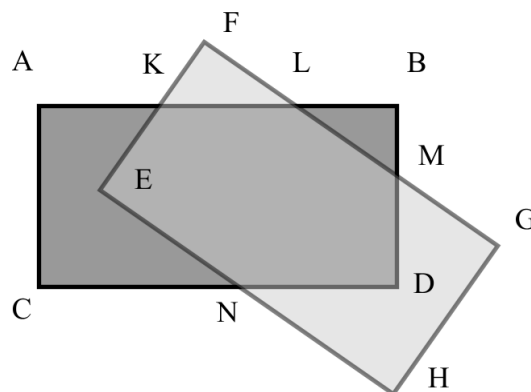
Теорема 1 [5]. Для произвольного выпуклого четырёхугольника на плоскости и $k > 0$ справедливо неравенство:

$$\frac{1}{(k+1)(k^2+k+1)} S \leq s \leq \frac{1}{2k^2+2k+1} S,$$

причём в случае $S > 0$ равенство $(k+1)(k^2+k+1)s = S$ выполняется лишь для четырёхугольников с двумя совпадающими вершинами, а любой четырёхугольник со свойством $(2k^2+2k+1)k = S$ помещается в непрерывное семейство четырёхугольников с тем же свойством, содержащее некоторый параллелограмм.

2. Задача Дж. Фике (J.W. Fickett)

Рассмотрим на евклидовой плоскости два конгруэнтных пересекающихся прямоугольника $P_1 = ABCD$ и $P_2 = EFGH$. Пусть L_1 – длина той части границы ∂P_1 первого прямоугольника, которая попадает во внутренность $\text{int}(P_2)$ второго. Аналогично, L_2 – длина части границы ∂P_2 второго прямоугольника, попадающей во внутренность $\text{int}(P_1)$ первого.



$$\frac{1}{3} L_1 \leq L_2 \leq 3L_1.$$

Задача Дж.В. Фике заключается в доказательстве неравенства.

Это неравенство оказалось доказано (в том числе) и при помощи компьютерных расчётов, точнее, справедливо.

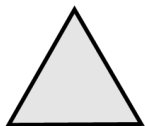
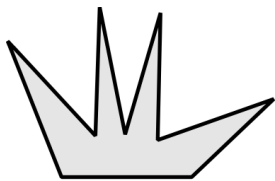
Теорема 2 [7]. Для пересекающихся конгруэнтных прямоугольников на евклидовой плоскости в обозначениях, приведённых выше, выполняется неравенство:

$$\frac{1}{3} L_1 \leq L_2 \leq 3L_1.$$

причём приведённое неравенство неулучшаемо.

3. Задача В.К. Ионина

Рассмотрим на евклидовой плоскости простую замкнутую ломаную, состоящую из нечётного числа звеньев длины 1. Требуется доказать, что площадь области, ограничиваемой данной ломаной, не меньше площади равностороннего треугольника со стороной 1, т.е. не меньше чем $\frac{\sqrt{3}}{4}$ (эта задача была сформулирована автору профессором В.К. Иониным).



Рассмотрим теперь несколько более общую постановку. Пусть на евклидовой плоскости дана простая замкнутая ломаная с m звеньями, одно из которых имеет длину $a \in [0,1]$, а длины остальных звеньев равны 1. Обозначим через S площадь области, ограничиваемой данной ломаной.

В работе [8] при помощи компьютерных расчётов, произведённых в Maple, доказана следующая теорема.

Теорема 3 [8, 11]. При $m = 2n-1$ справедливо неравенство $S \geq \frac{a}{4}\sqrt{4-a^2}$, при $m=2n$ справедливо неравенство $S \geq \frac{1-a}{4}\sqrt{4-(1-a)^2}$, причём оба эти неравенства неумлучшаемы.

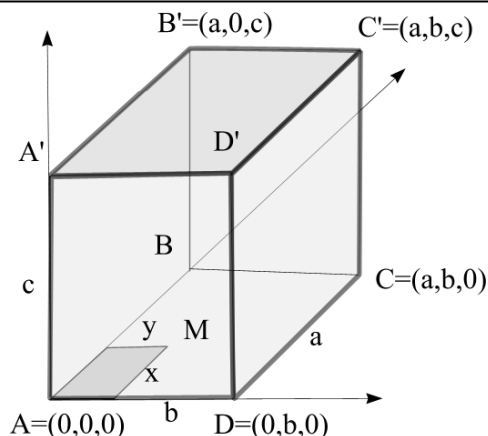
В качестве следствия при $a=1$ получается решение сформулированной выше задачи В.К. Ионина.

Теорема 4 [8, 11]. Любая замкнутая ломаная на евклидовой плоскости, состоящая из нечётного числа звеньев длины 1, ограничивает область с площадью не менее чем $\frac{\sqrt{3}}{4}$.

Следует отметить, что в работе венгерских математиков [11] теоремы 3 и 4 были существенно обобщены (работы [8, 11] были написаны независимо друг от друга и практически одновременно).

4. Задача о внутреннем расстоянии на поверхности параллелепипеда

Рассмотрим в трёхмерном евклидовом пространстве прямой параллелепипед $P=ABCD A'B'C'D'$ со сторонами длины $|AB|=a$, $|AD|=b$, $|AA'|=c$, $a \leq b \leq c$.



Объектом исследования является внутреннее расстояние между точками на поверхности $S = \partial P$ параллелепипеда P . Следует напомнить, что внутренним расстоянием $d(M, N)$ между точками $M \in S$ и $N \in S$ называется минимум длин ломаных, лежащих в S и соединяющих точки M и N . Достаточно сложной задачей является определение самой далёкой точки поверхности от заданной точки в смысле внутреннего расстояния. Даже в том случае, когда в качестве фиксированной точки берётся вершина параллелепипеда, не вполне ясно, какая точка будет от неё самой удалённой. Например, для куба самой далёкой точкой от вершины является противоположная вершина, в случае же параллелепипеда с размерами $1 \times 1 \times 2$ противоположная вершина не является самой удалённой точкой.

При помощи компьютерных расчётов, произведённых в Maple, был выявлен критерий существования точки, удалённой в смысле внутреннего расстояния от данной вершины более, чем противоположная вершина. Справедлива следующая теорема.

Теорема 5 [6]. Для прямого параллелепипеда в трёхмерном евклидовом пространстве с длинами сторон $a \leq b \leq c$ необходимым и достаточным условием существования точки на поверхности параллелепипеда, расположенной далее в смысле внутреннего расстояния от заданной вершины, чем противоположная вершина, является неравенство:

$$2c^2 - 2bc - ac - ab > 0.$$

Позже этот результат был получен другим методом в работе [1].

5. О геодезическом диаметре поверхности параллелепипеда

Под выпуклой поверхностью подразумевается граница ограниченного выпуклого тела в трёхмерном евклидовом пространстве. Вычисление геодезического диаметра выпуклых поверхностей является чрезвычайно сложной задачей, которая в настоящий момент решена лишь для поверхностей очень специального вида. До сих пор неизвестно никаких общих

методов для вычисления геодезического диаметра многогранных поверхностей. Необходимо отметить, что знание геодезического диаметра некоторых классов выпуклых поверхностей позволяет получить соответствующие оценки на экстремальные значения некоторых естественных функционалов, заданных на множестве выпуклых тел, при разнообразных ограничениях на геодезический диаметр. Поэтому вычисление геодезического диаметра (помимо чисто теоретического интереса) имеет и прикладное значение.

При помощи расчётов в Maple были найдены явные формулы для вычисления геодезического диаметра поверхности произвольного прямого параллелепипеда в трёхмерном евклидовом пространстве.

Рассмотрим в трёхмерном евклидовом пространстве прямой параллелепипед $ABCD A'B'C'D'$ со сторонами длины $|AB|=a$, $|AD|=b$, $|AA'|=c$, $a \leq b \leq c$ (см. рис. к задаче о внутреннем расстоянии на поверхности параллелепипеда). Объектом исследования является внутреннее расстояние между точками M и N на поверхности $S = \partial P$ параллелепипеда P или минимум длин ломаных, лежащих в S и соединяющих точки M и N . Неформально внутреннее расстояние есть длина кратчайшего пути, который должен преодолеть паук между двумя точками на границе комнаты (стены, пол, потолок). Геодезическим диаметром параллелепипеда (поверхности параллелепипеда) называется максимум внутренних расстояний между парами точек на поверхности параллелепипеда. Геодезический диаметр параллелепипеда P будем обозначать через $D(P)$.

Положим $M = \{(a, b, c) \in R^3, 0 < a \leq b \leq c\}$,

$$ME = \{(a, b, c) \in M \mid \sqrt{\max\{0, a^2 + 2ab - 2bc\}} + \sqrt{\max\{0, b^2 + 2ab - 2ac\}} \geq 2c - a - b\}.$$

При помощи компьютерных расчётов, произведённых в Maple, был выявлен следующий основной результат.

Теорема 6 [12]. Пусть $D(P)$ – геодезический диаметр поверхности прямого параллелепипеда P со сторонами длины $0 \leq a \leq b \leq c$.

Тогда, если $(a, b, c) \in ME$, то $D(P) = \sqrt{(a+b)^2 + c^2}$.

Если $(a, b, c) \in M \setminus ME$ и $a^2 b^2 \leq c^2(b-a)(a+b+2c)$, то

$$D(P) = \sqrt{b^2 + 3c^2 + 2b(a+c) - 2c\sqrt{(b+c)^2 - 2a(c-b) - a^2}}.$$

Если $(a, b, c) \in M \setminus ME$ и $a^2 b^2 \geq c^2(b-a)(a+b+2c)$, то

$D(P) = l$, где l – единственное вещественное решение уравнения:

$$\sqrt{l^2 - (a+c)^2} + \sqrt{l^2 - (b+c)^2} + \sqrt{2l^2 - (a+b+c)^2} = c$$

с условием $l \geq \max\{b+c, \sqrt{(a+b)^2 + c^2}\}$.

Результат теоремы 6 даёт возможность отыскания экстремальных значений различных функционалов, определённых на множестве прямоугольных параллелепипедов, с ограничением на геодезический диаметр. Например, естественной задачей такого рода является определение прямоугольного параллелепипеда с максимальной площадью среди параллелепипедов с заданным геодезическим диаметром.

Теорема 7 [12]. Среди прямоугольных параллелепипедов с заданным геодезическим диаметром максимальную площадь поверхности имеет параллелепипед с отношением рёбер $a:b:c=1:1:\sqrt{2}$. Другими словами, для всех прямоугольных параллелепипедов с рёбрами длины $0 \leq a \leq b \leq c$ выполнено неравенство:

$$ab + ac + bc \leq \frac{1 + 2\sqrt{2}}{6} (D(P))^2,$$

обращающееся в равенство тогда и только тогда, когда $a:b:c=1:1:\sqrt{2}$.

Полученные результаты могут быть использованы в дальнейших исследованиях по выпуклой геометрии и геометрии пространств с внутренней метрикой. В то же время эти результаты, являясь точными, могут служить для проверки эффективности алгоритмов и программных продуктов, предназначенных для вычисления стандартных характеристик геометрических объектов. Это важно с той точки зрения, что в настоящее время неизвестны оптимальные алгоритмы для вычисления таких важных в геометрии и приложениях величин, как внутреннее расстояние на поверхности заданного выпуклого тела, внутренний диаметр выпуклой поверхности и изопериметрические дефекты для плоских и пространственных областей.

Список литературы

1. Вялый М.Н. Кратчайшие пути на поверхности параллелепипеда // Математическое просвещение. – 2005. – Сер. 5. – № 9. – С. 203–206.
2. Говорухин В.Н., Цибулин В.Г. Компьютер в математическом исследовании: Maple, MATLAB, LaTeX. – СПб.: Питер. 2001. – 624 с.
3. Дьяконов В.П. Maple 9 в математике, физике и образовании. – М.: СОЛОН-Пресс, 2004. – 688 с.
4. Матросов А.В. Maple 6. Решение задач высшей математики и механики. – СПб.: БХВ-Петербург, 2001. – 528 с.
5. Никоноров Ю.Г., Никонорова Ю.В. Решение обобщённой задачи Поповичи // Труды Рубцовского индустриального института. – Рубцовск, 2000. – Т. 7. – С. 229–232.

6. Никоноров Ю.Г., Никонорова Ю.В. О внутренней геометрии поверхности прямоугольного параллелепипеда // Труды Рубцовского промышленного института. – Рубцовск, 2000. – Т. 7. – С. 222–228.
7. Никонорова Ю.В. Об одной экстремальной задаче на евклидовой плоскости // Математические труды. – 2001. – Т. 4. – № 1. – С. 111–121.
8. Никонорова Ю.В. Об одной изопериметрической задаче на евклидовой плоскости // Труды Рубцовского промышленного института. – Рубцовск, 2001. – Т. 9. – С. 66–72.
9. Никоноров Ю.Г., Никонорова Ю.В. Применение системы Maple к решению геометрических задач. – 2-е изд. – Рубцовск: Изд-во АлтГУ, 2005. – 80 с.
10. Шклярский Д.О., Ченцов Н.Н., Яглом И.М. Геометрические оценки и задачи из комбинаторной геометрии. – М.: Наука, 1974. – 384 с.
11. Böröczky K., Kertész G., Makai, Jr.E. The minimum area of a simple polygon with given side lengths // Periodica Mathematica Hungarica. – 1999. – No 39 (1–3). – P. 33–49.
12. Nikonorov Yu.G., Nikonorova Yu.V. The intrinsic diameter of the surface of a parallelepiped // Discrete Comput. Geom, 2008. – No 40 (4). – P. 504–527.

УДК 12.54:004.9

Е.Н. Коровин

ПРИМЕНЕНИЕ ПАКЕТА АНАЛИТИЧЕСКИХ ВЫЧИСЛЕНИЙ MAPLE ПРИ НАХОЖДЕНИИ МЕТРИКИ И КРИВИЗНЫ ДЛЯ 6-МЕРНЫХ НИЛЬПОТЕНТНЫХ СИМПЛЕКТИЧЕСКИХ ГРУПП ЛИ

В работе [1] прямым вычислением были получены многопараметрические семейства комплексных структур для разложимых 6-мерных нильпотентных групп Ли. В работе [2] приведена классификационная теорема, в которой предоставлен список всех 6-мерных симплектических нильпотентных групп Ли и в явном виде представлены симплектические формы для каждой группы.

В данной работе найдены левоинвариантные псевдокелеровы структуры на одной из групп Ли, исследованы свойства кривизны. Показано, что соответствующие метрики образуют пятипараметрическое семейство неплоских Риччи-плоских метрик.

Все вычисления проводятся по следующей схеме:

- вводятся все необходимые данные и формулы вычисления;
- проводится анализ полученных данных. Вводятся отношения между параметрами, которые следуют из условий задачи;
- анализ конечного результата.

Стоит отметить, что зависимость параметров в случае нильпотентных симплектических групп Ли, приведённых в [2], достаточно просты. При дальнейшем исследовании, когда базис приводится к ортогональному виду, вычисления заметно усложняются и использование пакета Maple становится менее рациональным.

1. Постановка задачи. Основные формулы вычисления

Будем рассматривать одну левоинвариантную 6-мерную нильпотентную симплектическую группу Ли из классификации [2]. В данной классификации рассматриваемая алгебра Ли имеет 13 номер. Для простоты будем обозначать её $A_{6,13}$. Структурные уравнения данной алгебры Ли записываются в следующем виде:

$$[X_1, X_2] = X_5, [X_1, X_3] = X_6, [X_2, X_4] = X_6, [X_4, X_3] = X_5.$$

В работе [1] найдены левоинвариантные комплексные структуры на шестимерной группе Ли $G_{6,13}$ с алгеброй Ли типа $A_{6,13}$. В работе [2] показано, что данная группа Ли является симплектической, и представители симплектической структуры задаются двумя формами вида:

$$\begin{aligned}\omega_1 &= \alpha_1 \wedge \alpha_6 + \alpha_2 \wedge \alpha_3 - \alpha_4 \wedge \alpha_5 \\ \omega_2 &= \alpha_1 \wedge \alpha_6 - \alpha_2 \wedge \alpha_3 + \alpha_4 \wedge \alpha_5.\end{aligned}$$

Поскольку рассматривается случай, когда все объекты левоинвариантные, достаточно будет проводить все вычисления на алгебре Ли данной группы.

Таким образом, алгебра Ли имеет всего три ненулевые структурные константы:

$$C_{12}^5 = 1, C_{13}^6 = 1, C_{24}^6 = 1, C_{43}^5 = 1.$$

Данная алгебра Ли имеет трёхмерный центр Z , образованный векторами E_4, E_5, E_6 . Общий элемент $X = x^1 E_1 + x^2 E_2 + x^3 E_3 + x^4 E_4 + x^5 E_5 + x^6 E_6$ алгебры Ли может быть представлен матрицей X порядка 6, первые четыре строки которой имеют вид $(0, x^1, x^5, x^6, 0, 0)$, $(0, 0, x^2, x^3, x^6, 0)$, $(0, 0, 0, 0, x^4, -x^5)$, $(0, 0, 0, 0, 0, x^4)$, а остальные строки – нулевые. Общий элемент соответствующей группы Ли $G_{6,13}$ имеет вид $g = Id + X$, где Id – единичная матрица.

Почти комплексная структура J является интегрируемой (комплексной) [3], если её тензор Нейенхейса $N(X, Y) = 2([JX, JY] - [X, Y] - J[X, JY] - J[JX, Y])$ обращается в нуль. В случае левоинвариантной почти комплексной структуры J на группе Ли тензор Нейенхейса легко выражается через структурные константы алгебры Ли:

$$N_{ij}^k = 2(J_i^l J_j^m C_{lm}^k - C_{ij}^k - J_m^k J_j^l C_{il}^m - J_m^k J_i^l C_{lj}^m).$$

Почти комплексная структура J называется ассоциированной с симплектической формой ω , если $\omega(JX, JY) = \omega(X, Y)$, для любых элементов $X, Y \in A_{6,13}$. В этом случае вторая форма

$$g_J(X, Y) = \omega(X, JY)$$

является симметричной и поэтому определяет на (псевдо)риманову метрику на алгебре Ли $A_{6,13}$. Если J – интегрируемая почти комплексная структура, то тройка (g_J, J, ω) определяет левоинвариантную (псевдо)келерову структуру на группе G_3 .

Как известно, компоненты римановой связности Γ_{ij}^k левоинвариантной метрики на группе Ли и тензора кривизны R_{ijk}^s выражаются через структурные константы. Для выбранного ранее базиса $\{E_1, E_2, E_3, E_4, E_5, E_6\}$ алгебры Ли $A_{6,13}$ имеем, $\nabla_{E_i} E_j = \Gamma_{ij}^k E_k$. Компоненты связности Γ_{ij}^k находим из шестичленной формулы [1], которая для левоинвариантных векторных полей X, Y, Z на группе Ли принимает вид: $2g(\nabla_X Y, Z) = g([X, Y], Z) + g([Z, X], Y) + g(X, [Z, Y])$.

Используя структурные константы C_{ij}^k в базисе $\{E_i\}$ алгебры Ли, получаем:

$$\Gamma_{ij}^n = \frac{1}{2} g^{kn} (C_{ij}^p g_{pk} + C_{ki}^p g_{pj} + C_{kj}^p g_{ip}).$$

Тензор кривизны определяется равенством:

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z,$$

где X, Y, Z – левоинвариантные векторные поля на группе Ли.

В базисе $\{E_i\}$ имеем:

$$\begin{aligned} R_{ijk}^s E_s &= R(E_i, E_j)E_k = \nabla_{E_i} \nabla_{E_j} E_k - \nabla_{E_j} \nabla_{E_i} E_k - \nabla_{[E_i, E_j]} E_k = \\ &= \Gamma_{ip}^s \Gamma_{jk}^p E_s - \Gamma_{jp}^s \Gamma_{ik}^p E_s - C_{ij}^p \Gamma_{pk}^s E_s, \\ R_{ijk}^s &= \Gamma_{ip}^s \Gamma_{jk}^p - \Gamma_{jp}^s \Gamma_{ik}^p - C_{ij}^p \Gamma_{pk}^s. \end{aligned}$$

Тензор Риччи определяется как свёртка тензора кривизны по первому и по четвёртому (верхнему) индексам: $Ric_{jk} = R_{ijk}^i$. Скалярная кривизна определяется как свёртка тензора Риччи $S = g^{jk} Ric_{jk}$. Будем рассматривать также (псевдо)риманов скалярный квадрат тензора кривизны:

$$NR = g^{ip} g^{jr} g^{ks} g_{lt} R_{ijk}^l R_{prs}^t.$$

Найдём ассоциированные с симплектической формой ω левоинвариантные комплексные структуры J на $A_{6,13}$, построим соответствующие метрики и вычислим их характеристики кривизны.

2. Полученные результаты

Для алгебры Ли $A_{6,13}$, заданной скобками Ли,

$$[X_1, X_2] = X_5, [X_1, X_3] = X_6, [X_2, X_4] = X_6, [X_4, X_3] = X_5.$$

Будут представлены все результаты в случае симплектической структуры:

$$\omega_1 = \alpha_1 \wedge \alpha_6 + \alpha_2 \wedge \alpha_3 - \alpha_4 \wedge \alpha_5$$

В этом случае имеем пятипараметрическое семейство комплексных структур и псевдоримановых метрик:

$$J_X = \begin{pmatrix} \xi_{11} & 0 & 0 & \xi_{14} & 0 & 0 \\ -\frac{\xi_{31}}{\xi_{14}\xi_{11}} & -J_{22} & J_{32} & J_{42} & 0 & 0 \\ \xi_{31} & -\frac{2\xi_{14}(1+\xi_{11}^2)}{1+\xi_{11}^2+\xi_{14}^2} & J_{22} & 0 & 0 & 0 \\ -\frac{1+\xi_{11}^2}{\xi_{14}} & 0 & 0 & -\xi_{11} & 0 & 0 \\ \xi_{51} & 0 & J_{42} & \xi_{54} & \xi_{11} & \xi_{14} \\ J_{16} & \xi_{31} & -\frac{\xi_{31}}{\xi_{14}\xi_{11}} & -\xi_{51} & -\frac{1+\xi_{11}^2}{\xi_{14}} & -\xi_{11} \end{pmatrix},$$

где

$$J_{22} = \frac{\xi_{11}(1+\xi_{11}^2-\xi_{14}^2)}{1+\xi_{11}^2+\xi_{14}^2}, J_{42} = \frac{\xi_{31}(1+\xi_{11}^2+\xi_{14}^2)}{2(1+\xi_{11}^2)}, J_{32} = \frac{1+2\xi_{11}^2+2\xi_{14}^2+\xi_{11}^4+\xi_{14}^4-2\xi_{11}^2\xi_{14}^2}{2(1+\xi_{11}^2+\xi_{14}^2)\xi_{14}},$$

$$J_{61} = \frac{2\xi_{54}\xi_{11}^4-4\xi_{51}\xi_{11}^3\xi_{14}+4\xi_{11}^2\xi_{54}-\xi_{14}\xi_{31}^2\xi_{11}^2-4\xi_{14}\xi_{51}\xi_{11}-\xi_{14}^3\xi_{31}^2+2\xi_{54}-\xi_{31}^2\xi_{14}}{2\xi_{14}^2(1+\xi_{11}^2)}.$$

$$g_X = \begin{pmatrix} J_{61} & \xi_{31} & -\frac{\xi_{31}}{\xi_{14}\xi_{11}} & -\xi_{51} & -\frac{1+\xi_{11}^2}{\xi_{14}} & -\xi_{11} \\ \xi_{31} & -\frac{2\xi_{14}(1+\xi_{11}^2)}{1+\xi_{11}^2+\xi_{14}^2} & J_{22} & 0 & 0 & 0 \\ -\frac{\xi_{31}}{\xi_{14}\xi_{11}} & J_{22} & J_{32} & J_{42} & 0 & 0 \\ -\xi_{51} & 0 & J_{42} & -\xi_{54} & -\xi_{11} & -\xi_{14} \\ -\frac{1+\xi_{11}^2}{\xi_{14}} & 0 & 0 & -\xi_{11} & 0 & 0 \\ -\xi_{11} & \xi_{31} & 0 & -\xi_{14} & 0 & 0 \end{pmatrix},$$

Тензор кривизны с точностью до антисимметрий имеет ровно одну ненулевую компоненту $R_{1414} = \frac{1+\xi_{11}^2+\xi_{14}^2}{\xi_{14}}$. Тензор Риччи нулевой при всех значениях параметра:

$$g_X = \begin{pmatrix} 0 & 0 & 0 & 0 & -\frac{1+\xi_{11}^2}{\xi_{14}} & -\xi_{11} \\ 0 & -\frac{2\xi_{14}(1+\xi_{11}^2)}{1+\xi_{11}^2+\xi_{14}^2} & J_{22} & 0 & 0 & 0 \\ 0 & J_{22} & J_{32} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\xi_{11} & -\xi_{14} \\ -\frac{1+\xi_{11}^2}{\xi_{14}} & 0 & 0 & -\xi_{11} & 0 & 0 \\ -\xi_{11} & 0 & 0 & -\xi_{14} & 0 & 0 \end{pmatrix},$$

Для простоты исследования занулим все параметры, не влияющие на тензор кривизны. В этом случае в псевдоримановой метрике останутся только компоненты, влияющие на кривизну, следующим шагом занулим параметр ξ_{11} .

Проведём замену базиса

$$Y_1 = X_2, Y_2 = X_3, Y_3 = X_5, Y_4 = X_1, Y_5 = X_6, Y_6 = X_4,$$

таким образом, получаем, что метрика в новом базисе имеет вид:

$$g_Y = \begin{pmatrix} -\frac{2\xi_{14}(1+\xi_{11}^2)}{1+\xi_{11}^2+\xi_{14}^2} & J_{22} & 0 & 0 & 0 & 0 \\ J_{22} & J_{32} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1+\xi_{11}^2}{\xi_{14}} & 0 & -\xi_{11} \\ 0 & 0 & -\frac{1+\xi_{11}^2}{\xi_{14}} & 0 & -\xi_{11} & 0 \\ 0 & 0 & 0 & -\xi_{11} & 0 & -\xi_{14} \\ 0 & 0 & -\xi_{11} & 0 & -\xi_{14} & 0 \end{pmatrix}.$$

Структурные уравнения перепишутся в виде:

$$[Y_4, Y_1] = Y_3, [Y_4, Y_2] = Y_5, [Y_1, Y_6] = Y_5, [Y_2, Y_6] = Y_3.$$

Тензор кривизны имеет с точностью до антисимметрий ровно 1 ненулевую компоненту R_{46} . Поскольку параметр ξ_{11} не влияет на количество ненулевых компонент тензора кривизны, предположим, что он равен нулю и метрика в этом случае упрощается до вида:

Поскольку векторы Y_3, Y_4, Y_5, Y_6 изотропны, а площадки Y_3, Y_4 и Y_5, Y_6 ортогональны, секционная кривизна второй плоскости Y_4, Y_6 не может быть определена в данном базисе:

$$g_Y = \begin{pmatrix} \frac{2\xi_{14}}{1+\xi_{14}^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1+\xi_{14}^2}{2\xi_{14}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\xi_{14}} & 0 & 0 \\ 0 & 0 & -\frac{1}{\xi_{14}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\xi_{14} \\ 0 & 0 & 0 & 0 & -\xi_{14} & 0 \end{pmatrix}.$$

Приведём базис к ортогональному виду с помощью следующей замены:

$$Z_1 = Y_1, Z_2 = Y_2, Z_3 = Y_3 + Y_4, Z_4 = Y_3 - Y_4, Z_5 = Y_5 + Y_6, Z_6 = Y_5 - Y_6.$$

Скобки Ли в новом базисе переписутся в следующем виде:

$$\begin{aligned} [Z_1, Z_3] &= -\frac{Z_3 + Z_4}{2}, [Z_1, Z_4] = \frac{Z_3 + Z_4}{2}, [Z_1, Z_5] = \frac{Z_5 + Z_6}{2}, \\ [Z_1, Z_6] &= -\frac{Z_5 + Z_6}{2}, [Z_2, Z_3] = -\frac{Z_5 + Z_6}{2}, [Z_2, Z_4] = \frac{Z_5 + Z_6}{2}, \\ [Z_2, Z_5] &= \frac{Z_3 + Z_4}{2}, [Z_2, Z_6] = -\frac{Z_3 + Z_4}{2}. \end{aligned}$$

Симплектическая структура в этом базисе принимает вид:

$$\omega_Z = \theta_1 \wedge \theta_2 + \frac{\theta_3 \wedge \theta_5}{2} - \frac{\theta_4 \wedge \theta_6}{2}.$$

Псевдориманова метрика принимает диагональный вид:

$$g_Y = \begin{pmatrix} \frac{2\xi_{14}}{1+\xi_{14}^2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1+\xi_{14}^2}{2\xi_{14}} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\sqrt{\frac{2}{|\xi_{14}|}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\frac{2}{|\xi_{14}|}} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{2|\xi_{14}|} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2|\xi_{14}|} \end{pmatrix},$$

таким образом получена псевдориманова метрика сигнатуры $(-, -, +, +, +, +)$. Секционная кривизна в новом базисе уже принимает ненулевые значения.

При переходе к общему виду пересчёт комплексной структуры и псевдоримановой метрике в данном базисе с помощью пакета Maple невозможен ввиду сильно расплывчатых ограничений на параметры.

3. Листинг программы вычислений

```
> restart;
> with(LinearAlgebra):with(tensor):with(linalg);
вводятся структурные константы
> C0:=array(sparse,1..6,1..6,1..6,[...]);
вводятся элементы симплектической структуры
> omega0:=array(sparse,1..6,1..6,[...]);
задаётся почти комплексная структура в общем виде:
> J:=array(1..6,1..6,[[xi11,xi12,xi13,xi14,xi15,xi16],
```

```
[xi21,xi22,xi23,xi24,xi25,xi26],[xi31,xi32,xi33,xi34,xi35,xi36],
[xi41,xi42,xi43,xi44,xi45,xi46],[xi51,xi52,xi53,xi54,xi55,xi56],
[xi61,xi62,xi63,xi64,xi65,xi66]]).
```

Проверка условия ассоциированности почти комплексной структуры симплектической формой $\omega(JX, JY) = \omega(X, Y)$:

```
> omega1:=array(1..6,1..6): for i1 to 6 do
> for j1 to 6 do
> k1:='k1':k2:='k2':
> omega1[i1,j1]:=sum(J[k1,i1]*omega0[k1,j1],'k1'=1..6)
+sum(omega0[i1,k2]*J[k2,j1],'k2'=1..6);
> od od ;
```

Тензор Нейенхейса:

```
> N:=array(1..6,1..6,1..6):for i to 6 do
> for j to 6 do
> for k to 6 do
> l:='l': m:='m':
> N[i,j,k]:=2*sum(sum(J[l,i]*J[m,j]*C0[l,m,k]-J[k,m]*J[l,j]*C0[i,l,m]
-J[k,m]*J[l,i]*C0[l,j,m],'l'=1..6),'m'=1..6)-2*C0[i,j,k];
> od od od ;
```

```
> i:='i': j:='j': k:='k':
```

условие $J^2 = -Id$:

```
> JJ:=simplify(multiply(J,J));
```

Выписывается псевдориманова метрика:

```
> GJ0:=simplify(multiply(omega0,J));
```

Матрица, обратная к метрике:

```
> GJinv0:=inverse(GJ0):
```

Вычисление компонент Γ^k_{ij} :

```
> Gamma0:=array(1..6,1..6,1..6):for i0 to 6 do
```

```
> for j0 to 6 do
```

```
> for p0 to 6 do
```

```
> k0:='k0': s0:='s0':
```

```
>
```

Gam-

```
ma0[i0,j0,p0]:=(1/2)*(sum(GJinv0[k0,p0]*sum(GJ0[s0,k0]*C0[i0,j0,s0]
+C0[k0,i0,s0]*GJ0[s0,j0]+GJ0[i0,s0]*C0[k0,j0,s0],'s0'=1..6),'k0'=1..6));
> od od od ;
```

Вычисление тензора кривизны:

```
> Riem0:=array(1..6,1..6,1..6,1..6):
```

```
> for ir0 to 6 do
```

```
> for jr0 to 6 do
```

```
> for kr0 to 6 do
```

```
> for sr0 to 6 do
```

```
> pr0:='pr0':
```

```

> Riem0[ir0,jr0,kr0,sr0]:=simplify(sum(Gamma0[ir0,pr0,sr0]*
Gamma0[jr0,kr0,pr0]-Gamma0[jr0,pr0,sr0]*Gamma0[ir0,kr0,pr0]-
C0[ir0,jr0,pr0]*Gamma0[pr0,kr0,sr0], 'pr0'=1..6));
> od od od od;
> i:='i': j:='j': k:='k': l:='l':
Перемещение верхнего индекса на четвертую позицию:
> Riem1:=array(1..6,1..6,1..6,1..6):for i1 to 6 do
>   for j1 to 6 do
>     for k1 to 6 do
>       for s1 to 6 do
>         p1:='p1':
> Riem1[i1,j1,k1,s1]:=simplify(sum(Riem0[i1,j1,k1,p1]*GJ0[p1,s1], 'p1'=1..6));
>       od od od od;
>     R1:=simplify(evalm(Riem1));
Вычисление тензора Риччи:
> RicP0:=array(1..6,1..6):for rn0 to 6 do
>   for rm0 to 6 do
>     ri0:='ri0':
>     RicP0[rm0,rm0]:=sum(Riem0[ri0,rm0,rm0,ri0], 'ri0'=1..6);
>   od od;

```

Список литературы

1. Magnin L. Technical report Complex Structures on Indecomposable 6-dimensional Nilpotent Real Lie Algebras // Preprint. – URL: <http://www.u-bourgogne.fr/monge/l.magnin>.
2. Khakimjanov Y., Goze M., Medina A. Symplectic or contact structures on Lie Groups // Preprint. – URL: <http://arxiv.org/abs/math.DG/0205290>.
3. Кобаяси Ш., Намидзу К. Основы дифференциальной геометрии: в 2 т. – М.: Наука, 1981.
4. MiatelloDottiI. Ricci curvature of left-invariant metrics on solvable unimodular Lie groups // Math. Zeit. – 1982. – Vol. 180. – P. 257–263.
5. Milnor J. Curvatures of left invariant metrics on Lie groups // Advances in Math. – 1976. – Vol. 21. – P. 293–329.
6. Морозов, В.В. Классификация нильпотентных алгебр Ли шестого порядка // Изв. вузов. Сер. Математика. – 1958. – № 4. – С. 161–174.

УДК 512.5

А.В. Коротков

МНОГОМЕРНЫЕ АЛГЕБРЫ ПОЛЕЙ СРАВНЕНИЙ ПО MOD=2

Статья посвящена основным принципам построения многомерных алгебр полей сравнений по mod = 2.

Ключевые слова: алгебра одномерных полей сравнения по mod = 2, алгебра двумерных полей сравнений по mod=2, алгебра n-мерных полей сравнений по mod = 2.

Алгебра одномерных полей сравнений по mod=2

Алгеброй одномерных полей сравнений по mod=2 назовём класс S объектов $\mathbf{a}=(a_1)$, $\mathbf{b}=(b_1)$, $\mathbf{c}=(c_1)$, ..., в котором определены две бинарные операции, обозначаемые как (логические) сложение и умножение, со следующими свойствами:

операция сложения

+	0	1
0	0	1
1	1	0

операция умножения

*	0	1
0	0	0
1	0	1

Для всех $\mathbf{a}=(a_1)$, $\mathbf{b}=(b_1)$, $\mathbf{c}=(c_1)$, ... из S имеют место:

1) (замкнутость) S содержит:

$$(\mathbf{a}_1)+(\mathbf{b}_1)=(\mathbf{a}_1+\mathbf{b}_1);$$

$$(\mathbf{a}_1)\cdot(\mathbf{b}_1)=(\mathbf{a}_1\cdot\mathbf{b}_1);$$

2) коммутативные законы:

$$(\mathbf{a}_1)+(\mathbf{b}_1)=(\mathbf{b}_1)+(\mathbf{a}_1), \text{ т.е. } \mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a};$$

$$(\mathbf{a}_1)\cdot(\mathbf{b}_1)=(\mathbf{b}_1)\cdot(\mathbf{a}_1), \text{ т.е. } \mathbf{a}\cdot\mathbf{b}=\mathbf{b}\cdot\mathbf{a};$$

3) ассоциативные законы:

$$(\mathbf{a}_1)+((\mathbf{b}_1)+(\mathbf{c}_1))=((\mathbf{a}_1)+(\mathbf{b}_1))+(\mathbf{c}_1),$$

$$\text{т.е. } \mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c};$$

$$(\mathbf{a}_1)\cdot((\mathbf{b}_1)\cdot(\mathbf{c}_1))=((\mathbf{a}_1)\cdot(\mathbf{b}_1))\cdot(\mathbf{c}_1),$$

$$\text{т.е. } \mathbf{a}\cdot(\mathbf{b}\cdot\mathbf{c})=(\mathbf{a}\cdot\mathbf{b})\cdot\mathbf{c};$$

4) дистрибутивные законы:

$$(\mathbf{a}_1)\cdot((\mathbf{b}_1)+(\mathbf{c}_1))=(\mathbf{a}_1)\cdot(\mathbf{b}_1)+(\mathbf{a}_1)\cdot(\mathbf{c}_1),$$

$$\text{т.е. } \mathbf{a}\cdot(\mathbf{b}+\mathbf{c})=(\mathbf{a}\cdot\mathbf{b})+(\mathbf{a}\cdot\mathbf{c});$$

5) (свойства идемпотентности):

$$(\mathbf{a}_1)=0, \text{ т.е. } \mathbf{a}+\mathbf{a}=\mathbf{0},$$

$$(\mathbf{a}_1)\cdot((\mathbf{a}_1)+(\mathbf{a}_1))=\mathbf{a}_1, \text{ т.е. } \mathbf{a}\cdot\mathbf{a}=\mathbf{a};$$

6) (свойства совместимости) – (не используются)

7) S содержит элементы $\mathbf{1}=(1)$, $\mathbf{0}=(0)$ такие, что для всякого элемента $\mathbf{a}=(a_1)$ из S

$$(a_1)+(0)=(a_1), \text{ т.е. } \mathbf{a}+\mathbf{0}=\mathbf{a},$$

$$(a_1)\cdot(1)=(a_1), \text{ т.е. } \mathbf{a}\cdot\mathbf{1}=\mathbf{a},$$

$$(a_1)\cdot(0)=(0), \text{ т.е. } \mathbf{a}\cdot\mathbf{0}=\mathbf{0},$$

$$(a_1)+(1)=(\bar{a}_1), \text{ т.е. } \mathbf{a}+\mathbf{1}=\bar{\mathbf{a}};$$

8) для каждого элемента $\mathbf{a}=(a_1)$ класс S содержит элемент $\bar{\mathbf{a}}=(\bar{a}_1)$ (дополнение элемента $\mathbf{a}=(a_1)$) такой, что

$$(a_1)+(\bar{a}_1)=1, \text{ т.е. } \mathbf{a}+\bar{\mathbf{a}}=\mathbf{1},$$

$$(a_1)\cdot(\bar{a}_1)=0, \text{ т.е. } \mathbf{a}\cdot\bar{\mathbf{a}}=\mathbf{0}.$$

В алгебре одномерных полей сравнений имеют место:

9) законы поглощения:

$$(a_1)\cdot((a_1)+(b_1))=(a_1)+(a_1)\cdot(b_1)=(a_1)\cdot((1)+(b_1))=(a_1)\cdot(\bar{b}_1), \text{ т.е. } \mathbf{a}\cdot(\mathbf{a}+\mathbf{b})=\mathbf{a}\cdot\bar{\mathbf{b}};$$

$$(a_1)+((a_1)\cdot(b_1))=(a_1)\cdot((1)+(b_1))=(a_1)\cdot(\bar{b}_1), \text{ т.е. } \mathbf{a}+\mathbf{a}\cdot\mathbf{b}=\mathbf{a}\cdot\bar{\mathbf{b}};$$

10) двойственность (законы де Моргана):

$$\overline{(a_1)+(b_1)}=(\bar{a}_1)\cdot(\bar{b}_1)+(a_1)\cdot(b_1), \text{ т.е. } \overline{\mathbf{a}+\mathbf{b}}=\bar{\mathbf{a}}\cdot\bar{\mathbf{b}}+\mathbf{a}\cdot\mathbf{b};$$

$$\overline{(a_1)\cdot(b_1)}=(\bar{a}_1)+(a_1)\cdot(\bar{b}_1), \text{ т.е. } \overline{\mathbf{a}\cdot\mathbf{b}}=\bar{\mathbf{a}}+\mathbf{a}\cdot\bar{\mathbf{b}};$$

$$11) \overline{(\bar{a}_1)}=(a_1), \text{ т.е. } \overline{\bar{\mathbf{a}}}=\mathbf{a},$$

$$(\bar{1})=(0), \text{ т.е. } \bar{\mathbf{1}}=\mathbf{0},$$

$$(\bar{0})=(1), \text{ т.е. } \bar{\mathbf{0}}=\mathbf{1};$$

$$12) (a_1)\cdot(\bar{b}_1)+(\bar{a}_1)\cdot(b_1)=(a_1)+(b_1), \text{ т.е. } \mathbf{a}\cdot\bar{\mathbf{b}}+\bar{\mathbf{a}}\cdot\mathbf{b}=\mathbf{a}+\mathbf{b};$$

$$(a_1)\cdot((\bar{a}_1)+(b_1))=(a_1)\cdot(b_1), \text{ т.е. } \mathbf{a}\cdot(\bar{\mathbf{a}}+\mathbf{b})=\mathbf{a}\cdot\mathbf{b}.$$

Приведём таблицу истинности для полей вычетов по модулю 2. В таблице две переменные $\mathbf{a}$ и $\mathbf{b}$ с двумя состояниями 0 и 1 образуют четыре комбинации состояний и зависящие от них 16 функций двух переменных. Выполнение операций подтверждается таблицей истинности.

Алгебра двумерных полей сравнений по mod=2

Алгеброй двумерных полей сравнений по mod=2 назовём класс S объектов $\mathbf{a}=(a_1, a_2)$, $\mathbf{b}=(b_1, b_2)$, $\mathbf{c}=(c_1, c_2)$, ..., в котором определены две бинарные операции, обозначаемые как (логические) сложение и умножение, со следующими свойствами:

для всех $\mathbf{a}=(a_1, a_2)$, $\mathbf{b}=(b_1, b_2)$, $\mathbf{c}=(c_1, c_2)$ из S имеют место:

1) (замкнутость) S содержит:

$$(a_1, a_2)+(b_1, b_2)=(a_1+b_1, a_2+b_2)=\mathbf{a}+\mathbf{b};$$

$$(a_1, a_2)\cdot(b_1, b_2)=(a_1\cdot b_1, a_2\cdot b_2)=\mathbf{a}\cdot\mathbf{b};$$

2) коммутативные законы:

$$\begin{aligned}(a_1, a_2) + (b_1, b_2) &= (a_1 + b_1, a_2 + b_2), \\ (b_1, b_2) + (a_1, a_2) &= (b_1 + a_1, b_2 + a_2), \text{ т.е. } \mathbf{a+b=b+a}; \\ (a_1, a_2) \cdot (b_1, b_2) &= (a_1 \cdot b_1, a_2 \cdot b_2), \\ (b_1, b_2) \cdot (a_1, a_2) &= (b_1 \cdot a_1, b_2 \cdot a_2), \text{ т.е. } \mathbf{a \cdot b = b \cdot a};\end{aligned}$$

3) ассоциативные законы:

$$\begin{aligned}(a_1, a_2) + ((b_1, b_2) + (c_1, c_2)) &= (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2)), \\ ((a_1, a_2) + (b_1, b_2)) + (c_1, c_2) &= ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2), \\ \text{т.е. } \mathbf{a+(b+c)=(a+b)+c}; \\ (a_1, a_2) \cdot ((b_1, b_2) \cdot (c_1, c_2)) &= (a_1 \cdot (b_1 \cdot c_1), a_2 \cdot (b_2 \cdot c_2)), \\ ((a_1, a_2) \cdot (b_1, b_2)) \cdot (c_1, c_2) &= ((a_1 \cdot b_1) \cdot c_1, (a_2 \cdot b_2) \cdot c_2), \\ \text{т.е. } \mathbf{a \cdot (b \cdot c) = (a \cdot b) \cdot c};\end{aligned}$$

4) дистрибутивные законы:

$$\begin{aligned}(a_1, a_2) \cdot ((b_1, b_2) + (c_1, c_2)) &= (a_1 \cdot (b_1 + c_1), a_2 \cdot (b_2 + c_2)), \\ (a_1, a_2) \cdot (b_1, b_2) + (a_1, a_2) \cdot (c_1, c_2) &= (a_1 \cdot b_1 + a_1 \cdot c_1, a_2 \cdot b_2 + a_2 \cdot c_2), \\ \text{т.е. } \mathbf{a \cdot (b+c) = (a \cdot b) + (a \cdot c)};\end{aligned}$$

5) свойства идемпотентности:

$$\begin{aligned}(a_1, a_2) + (a_1, a_2) &= (0, 0), \text{ т.е. } \mathbf{a+a=0}, \\ (a_1, a_2) \cdot (a_1, a_2) &= (a_1, a_2), \text{ т.е. } \mathbf{a \cdot a=a};\end{aligned}$$

6) (свойства совместимости) – (не используются);

7) S содержит элементы $\mathbf{1}=(1, 1)$ и $\mathbf{0}=(0, 0)$ такие, что для всякого элемента $\mathbf{a}=(a_1, a_2)$ из S :

$$\begin{aligned}(a_1, a_2) + (0, 0) &= (a_1, a_2), \text{ т.е. } \mathbf{a+0=a}, \\ (a_1, a_2) \cdot (1, 1) &= (a_1, a_2), \text{ т.е. } \mathbf{a \cdot 1=a}, \\ (a_1, a_2) \cdot (0, 0) &= (0, 0), \text{ т.е. } \mathbf{a \cdot 0=0}, \\ (a_1, a_2) + (1, 1) &= (\bar{a}_1, \bar{a}_2), \text{ т.е. } \mathbf{a+1=\bar{a}};\end{aligned}$$

8) для каждого элемента $\mathbf{a}=(a_1, a_2)$ класс S содержит элемент $\bar{\mathbf{a}}=(\bar{a}_1, \bar{a}_2)$ (дополнение элемента $\mathbf{a}=(a_1, a_2)$) такой, что

$$\begin{aligned}(a_1, a_2) + (\bar{a}_1, \bar{a}_2) &= (a_1 + \bar{a}_1, a_2 + \bar{a}_2) = (1, 1), \text{ т.е. } \mathbf{a+\bar{a}=1}, \\ (a_1, a_2) \cdot (\bar{a}_1, \bar{a}_2) &= (a_1 \cdot \bar{a}_1, a_2 \cdot \bar{a}_2) = (0, 0), \text{ т.е. } \mathbf{a \cdot \bar{a}=0}.\end{aligned}$$

В каждой алгебре двумерных полей сравнений по $\text{mod}=2$ имеют место:

9) законы поглощения:

$$\begin{aligned}(a_1, a_2) \cdot ((a_1, a_2) + (b_1, b_2)) &= (a_1 \cdot (a_1 + b_1), a_2 \cdot (a_2 + b_2)) = (a_1 \cdot \bar{b}_1, a_2 \cdot \bar{b}_2), \\ \text{т.е. } \mathbf{a \cdot (a+b) = a \cdot \bar{b}}, \\ (a_1, a_2) + ((a_1, a_2) \cdot (b_1, b_2)) &= (a_1 + (a_1 \cdot b_1), a_2 + (a_2 \cdot b_2)) = (a_1 \cdot \bar{b}_1, a_2 \cdot \bar{b}_2), \\ \text{т.е. } \mathbf{a+a \cdot b = a \cdot \bar{b}};\end{aligned}$$

10) двойственность (законы де Моргана):

$$\begin{aligned} & \overline{(a_1, a_2) + (b_1, b_2)} = \overline{(a_1 + b_1, a_2 + b_2)} = \\ & = (\overline{a_1 + b_1}, \overline{a_2 + b_2}) = (\overline{a_1}, \overline{a_2}) \cdot (\overline{b_1}, \overline{b_2}) = (a_1, a_2) \cdot (b_1, b_2), \\ & \text{т.е. } \mathbf{a + b = a \cdot b + a \cdot b}, \\ & \overline{(a_1, a_2) \cdot (b_1, b_2)} = \overline{(a_1 \cdot b_1, a_2 \cdot b_2)} = \\ & = (\overline{a_1 \cdot b_1}, \overline{a_2 \cdot b_2}) = (\overline{a_1}, \overline{a_2}) + (\overline{b_1}, \overline{b_2}), \\ & \text{т.е. } \mathbf{a \cdot b = a + a \cdot b}; \end{aligned}$$

11) $\overline{\overline{(a_1, a_2)}} = \overline{(\overline{a_1}, \overline{a_2})} = (a_1, a_2)$, т.е. $\overline{\overline{\mathbf{a}}} = \mathbf{a}$,

$$(\overline{1}, \overline{1}) = (\overline{1}, \overline{1}) = (0, 0), \text{ т.е. } \overline{\mathbf{1}} = \mathbf{0},$$

$$(\overline{0}, \overline{0}) = (\overline{0}, \overline{0}) = (1, 1), \text{ т.е. } \overline{\mathbf{0}} = \mathbf{1};$$

12) $(a_1, a_2)(\overline{b_1}, \overline{b_2}) + (\overline{a_1}, \overline{a_2})(b_1, b_2) = (a_1 \overline{b_1} + \overline{a_1} b_1, a_2 \overline{b_2} + \overline{a_2} b_2) =$
 $= (a_1 + b_1, a_2 + b_2) = (a_1, a_2) + (b_1, b_2)$, т.е. $\mathbf{a \overline{b} + \overline{a} b = a + b}$,
 $(a_1, a_2) \cdot ((\overline{a_1}, \overline{a_2}) + (b_1, b_2)) = (a_1 \cdot (\overline{a_1} + b_1), a_2 \cdot (\overline{a_2} + b_2)) =$
 $= (a_1 \cdot b_1, a_2 \cdot b_2) = (a_1, a_2) \cdot (b_1, b_2)$, т.е. $\mathbf{a \cdot (\overline{a} + b) = a \cdot b}$.

Алгебра n-мерных полей сравнений по mod=2

Алгеброй n-мерных полей сравнений по mod=2 назовём класс S объектов $\mathbf{a}=(a_1, a_2, \dots, a_n)$, $\mathbf{b}=(b_1, b_2, \dots, b_n)$, $\mathbf{c}=(c_1, c_2, \dots, c_n)$, ..., в котором определены две бинарные операции, обозначаемые как (логические) сложение и умножение, со следующими свойствами:

для всех $\mathbf{a}=(a_1, a_2, \dots, a_n)$, $\mathbf{b}=(b_1, b_2, \dots, b_n)$, $\mathbf{c}=(c_1, c_2, \dots, c_n)$, ... из S имеют место:

1) (замкнутость) S содержит:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = \mathbf{a + b};$$

$$(a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n) = \mathbf{a \cdot b};$$

2) коммутативные законы:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n),$$

$$(b_1, b_2, \dots, b_n) + (a_1, a_2, \dots, a_n) = (b_1 + a_1, b_2 + a_2, \dots, b_n + a_n), \text{ т.е. } \mathbf{a + b = b + a},$$

$$(a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n),$$

$$(b_1, b_2, \dots, b_n) \cdot (a_1, a_2, \dots, a_n) = (b_1 \cdot a_1, b_2 \cdot a_2, \dots, b_n \cdot a_n), \text{ т.е. } \mathbf{a \cdot b = b \cdot a};$$

3) ассоциативные законы:

$$(a_1, a_2, \dots, a_n) + ((b_1, b_2, \dots, b_n) + (c_1, c_2, \dots, c_n)) = (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2), \dots, a_n + (b_n + c_n)),$$

$$((a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)) + (c_1, c_2, \dots, c_n) = ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2, \dots, (a_n + b_n) + c_n), \text{ т.е. } \mathbf{a + (b + c) = (a + b) + c},$$

$$(a_1, a_2, \dots, a_n) \cdot ((b_1, b_2, \dots, b_n) \cdot (c_1, c_2, \dots, c_n)) =$$

$$= (a_1 \cdot (b_1 \cdot c_1), a_2 \cdot (b_2 \cdot c_2), \dots, a_n \cdot (b_n \cdot c_n)),$$

$$((a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n)) \cdot (c_1, c_2, \dots, c_n) =$$

$$= ((a_1 \cdot b_1) \cdot c_1, (a_2 \cdot b_2) \cdot c_2, \dots, (a_n \cdot b_n) \cdot c_n), \text{ т.е. } \mathbf{a \cdot (b \cdot c) = (a \cdot b) \cdot c};$$

4) дистрибутивные законы:

$$\begin{aligned} (a_1, a_2, \dots, a_n) \cdot ((b_1, b_2, \dots, b_n) + (c_1, c_2, \dots, c_n)) &= \\ = (a_1 \cdot (b_1 + c_1), a_2 \cdot (b_2 + c_2), \dots, a_n \cdot (b_n + c_n)), \\ (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) + (a_1, a_2, \dots, a_n) \cdot (c_1, c_2, \dots, c_n) &= \\ = (a_1 \cdot b_1 + a_1 \cdot c_1, a_2 \cdot b_2 + a_2 \cdot c_2, \dots, a_n \cdot b_n + a_n \cdot c_n), \text{ т.е. } \mathbf{a \cdot (b+c) = (a \cdot b) + (a \cdot c)}; \end{aligned}$$

5) свойства идемпотентности:

$$\begin{aligned} (a_1, a_2, \dots, a_n) + (a_1, a_2, \dots, a_n) &= (0, 0, \dots, 0), \text{ т.е. } \mathbf{a+a=0}, \\ (a_1, a_2, \dots, a_n) \cdot (a_1, a_2, \dots, a_n) &= (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a \cdot a = a}; \end{aligned}$$

6) свойства совместимости – не используются;

7) S содержит элементы $\mathbf{1}=(1,1,\dots,1)$ и $\mathbf{0}=(0,0,\dots,0)$ такие, что для всякого элемента $\mathbf{a}=(a_1, a_2, \dots, a_n)$ из S:

$$\begin{aligned} (a_1, a_2, \dots, a_n) + (0, 0, \dots, 0) &= (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a+0=a}, \\ (a_1, a_2, \dots, a_n) \cdot (1, 1, \dots, 1) &= (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a \cdot 1 = a}, \\ (a_1, a_2, \dots, a_n) \cdot (0, 0, \dots, 0) &= (0, 0, \dots, 0), \text{ т.е. } \mathbf{a \cdot 0 = 0}, \\ (a_1, a_2, \dots, a_n) + (1, 1, \dots, 1) &= (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n), \text{ т.е. } \mathbf{a+1=\bar{a}}; \end{aligned}$$

8) для каждого элемента $\mathbf{a}=(a_1, a_2, \dots, a_n)$ класс S содержит элемент $\mathbf{\bar{a}}=(\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n)$ (дополнение элемента $\mathbf{a}=(a_1, a_2, \dots, a_n)$) такой, что

$$\begin{aligned} (a_1, a_2, \dots, a_n) + (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) &= (a_1 + \bar{a}_1, a_2 + \bar{a}_2, \dots, a_n + \bar{a}_n) = (1, 1, \dots, 1), \\ \text{т.е. } \mathbf{a + \bar{a} = 1}, \\ (a_1, a_2, \dots, a_n) \cdot (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) &= (a_1 \cdot \bar{a}_1, a_2 \cdot \bar{a}_2, \dots, a_n \cdot \bar{a}_n) = (0, 0, \dots, 0), \\ \text{т.е. } \mathbf{a \cdot \bar{a} = 0}. \end{aligned}$$

В каждой алгебре n-мерных полей сравнений по mod=2 имеют место:

9) законы поглощения:

$$\begin{aligned} (a_1, a_2, \dots, a_n) \cdot ((a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)) &= \\ = (a_1 \cdot (a_1 + b_1), a_2 \cdot (a_2 + b_2), \dots, a_n \cdot (a_n + b_n)) &= (a_1 \bar{b}_1, a_2 \bar{b}_2, \dots, a_n \bar{b}_n), \\ \text{т.е. } \mathbf{a \cdot (a+b) = a \cdot \bar{b}}, \\ (a_1, a_2, \dots, a_n) + ((a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n)) &= \\ = (a_1 + (a_1 \cdot b_1), a_2 + (a_2 \cdot b_2), \dots, a_n + (a_n \cdot b_n)) &= (a_1 \bar{b}_1, a_2 \bar{b}_2, \dots, a_n \bar{b}_n), \\ \text{т.е. } \mathbf{a + a \cdot b = a \cdot \bar{b}}; \end{aligned}$$

10) двойственность (законы де Моргана):

$$\begin{aligned} \overline{(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)} &= \overline{(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)} = \\ = (\bar{a}_1 \cdot \bar{b}_1, \bar{a}_2 \cdot \bar{b}_2, \dots, \bar{a}_n \cdot \bar{b}_n) &= \\ = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) \cdot (\bar{b}_1, \bar{b}_2, \dots, \bar{b}_n) &+ (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n), \\ \text{т.е. } \mathbf{\overline{a + b} = \bar{a} \cdot \bar{b} + a \cdot b}, \\ \overline{(a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n)} &= \overline{(a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n)} = \\ = (\bar{a}_1 + \bar{a}_1 \bar{b}_1, \bar{a}_2 + \bar{a}_2 \bar{b}_2, \dots, \bar{a}_n + \bar{a}_n \bar{b}_n) &= (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) + (\bar{b}_1, \bar{b}_2, \dots, \bar{b}_n), \\ \text{т.е. } \mathbf{\overline{a \cdot b} = \bar{a} + \bar{b}}; \end{aligned}$$

- 11) $\overline{\overline{(a_1, a_2, \dots, a_n)}} = \overline{\overline{(a_1, a_2, \dots, a_n)}} = (a_1, a_2, \dots, a_n)$, т.е. $\overline{\overline{\mathbf{a}}} = \mathbf{a}$,
 $\overline{(1, 1, \dots, 1)} = (\overline{1}, \overline{1}, \dots, \overline{1}) = (0, 0, \dots, 0)$, т.е. $\overline{\mathbf{1}} = \mathbf{0}$,
 $\overline{(0, 0, \dots, 0)} = (\overline{0}, \overline{0}, \dots, \overline{0}) = (1, 1, \dots, 1)$, т.е. $\overline{\mathbf{0}} = \mathbf{1}$;
- 12) $(a_1, a_2, \dots, a_n)(\overline{b_1}, \overline{b_2}) + (\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) \cdot (b_1, b_2, \dots, b_n) =$
 $= (a_1 \overline{b_1} + \overline{a_1} \cdot b_1, a_2 \overline{b_2} + \overline{a_2} \cdot b_2, \dots, a_n \overline{b_n} + \overline{a_n} \cdot b_n) =$
 $= (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)$,
т.е. $\mathbf{a} \cdot \overline{\mathbf{b}} + \overline{\mathbf{a}} \cdot \mathbf{b} = \mathbf{a} + \mathbf{b}$,
 $(a_1, a_2, \dots, a_n) \cdot ((\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) + (b_1, b_2, \dots, b_n)) =$
 $= (a_1 \cdot (\overline{a_1} + b_1), a_2 \cdot (\overline{a_2} + b_2), \dots, a_n \cdot (\overline{a_n} + b_n)) =$
 $= (a_1 \cdot \overline{a_1}, a_2 \cdot \overline{a_2}, \dots, a_n \cdot \overline{a_n}) = (\overline{a_1}, \overline{a_2}, \dots, \overline{a_n})$,
т.е. $\mathbf{a} \cdot (\overline{\mathbf{a}} + \mathbf{b}) = \overline{\mathbf{a}}$.

Таким образом, свойства алгебр многомерных полей сравнений по mod=2 повторяют свойства алгебры одномерных полей сравнений по mod=2. Вместе с тем, эти алгебры создают прецедент по отношению к булевым алгебрам в связи с изменением операции логического сложения, а следовательно, требуют расширения понятий основ кибернетики, алгебры классов, теории множеств и топологии по отношению к операциям логического сложения (объединения), логического умножения (пересечения) и взятия дополнения. То же самое можно сказать о сферах технического, физического, биологического и других сфер применения отмеченных алгебр.

a	0	0	1	1
b	0	1	0	1
0	0	0	0	0
a·b	0	0	0	1
a· $\overline{b}$	0	0	1	0
a	0	0	1	1
$\overline{a} \cdot b$	0	1	0	0
b	0	1	0	1
a+b	0	1	1	0
$\overline{\overline{a \cdot b}}$	0	1	1	1
$\overline{a \cdot b}$	1	0	0	0
$\overline{a + b}$	1	0	0	1
$\overline{b}$	1	0	1	0
$\overline{\overline{a \cdot b}}$	1	0	1	1
$\overline{a}$	1	1	0	0

Окончание табл.

$\overline{a \cdot b}$	1	1	0	1
$\overline{a \cdot b}$	1	1	1	0
1	1	1	1	1
$a+a$	0	0	0	0
$a \cdot a$	0	0	1	1
$a+1$	1	1	0	0
$a+\overline{a}$	1	1	1	1
$a \cdot \overline{a}$	0	0	0	0
$a \cdot (a+b)$	0	0	0	1
$a+a \cdot b$	0	0	0	1
$\overline{a \cdot b} + a \cdot b$	1	0	0	1
$\overline{a} + a \cdot \overline{b}$	1	1	1	0
$a \cdot \overline{b} + \overline{a} \cdot b$	0	1	1	0
$\overline{a} + b$	1	0	0	1
$a \cdot (\overline{a} + b)$	0	0	0	1

Список литературы

1. Коротков А.В. Небулевы алгебры логики // Информационные системы и технологии. Теория и практика: сб. науч. тр. / под ред. А.Н. Берёзы. – Шахты: Изд-во ЮРГУЭС, 2008. – 188 с.

УДК 512.5

А.В. Коротков

МНОГОМЕРНЫЕ БУЛЕВЫ АЛГЕБРЫ

Статья посвящена основным принципам построения многомерных булевых алгебр. Показано, что свойства многомерных булевых алгебр повторяют свойства одномерной булевой алгебры.

Ключевые слова: одномерная булева алгебра, двухмерная булева алгебра, n-мерная булева алгебра.

Одномерные Булевы алгебры

Одномерной Булевой алгеброй называется [1] класс S объектов $a=a_1$, $b=b_1$, $c=c_1$, ..., в котором определены две бинарные операции, обозначаемые как (логические) сложение и умножение, со следующими свойствами:

для всех $a=a_1$, $b=b_1$, $c=c_1$, ... из S имеют место:

1) (замкнутость) S содержит:

$$a_1+b_1=a+b;$$

$$a_1 \cdot b_1=a \cdot b;$$

2) коммутативные законы:

$$a_1 + b_1 = b_1 + a_1, \text{ т.е. } \mathbf{a+b=b+a};$$

$$a_1 \cdot b_1 = b_1 \cdot a_1, \text{ т.е. } \mathbf{a \cdot b = b \cdot a};$$

3) ассоциативные законы:

$$a_1 + (b_1 + c_1) = (a_1 + b_1) + c_1, \text{ т.е. } \mathbf{a+(b+c)=(a+b)+c};$$

$$a_1 \cdot (b_1 \cdot c_1) = (a_1 \cdot b_1) \cdot c_1, \text{ т.е. } \mathbf{a \cdot (b \cdot c) = (a \cdot b) \cdot c};$$

4) дистрибутивные законы:

$$a_1 \cdot (b_1 + c_1) = a_1 \cdot b_1 + a_1 \cdot c_1, \text{ т.е. } \mathbf{a \cdot (b+c) = (a \cdot b) + (a \cdot c)};$$

$$a_1 + (b_1 \cdot c_1) = (a_1 + b_1) \cdot (a_1 + c_1), \text{ т.е. } \mathbf{a + (b \cdot c) = (a+b) \cdot (a+c)};$$

5) (свойства идемпотентности):

$$a_1 + a_1 = a_1, \text{ т.е. } \mathbf{a+a=a};$$

$$a_1 \cdot a_1 = a_1, \text{ т.е. } \mathbf{a \cdot a = a};$$

6) (свойства совместимости):

$$a_1 + b_1 = \begin{cases} a_1, & \text{если } a_1 \cdot b_1 = b_1, \\ b_1, & \text{если } a_1 \cdot b_1 = a_1, \end{cases}$$

$$\text{т.е. } \mathbf{a+b} = \begin{cases} \mathbf{a}, & \text{если } \mathbf{a \cdot b = b}; \\ \mathbf{b}, & \text{если } \mathbf{a \cdot b = a}; \end{cases}$$

$$a_1 \cdot b_1 = \begin{cases} a_1, & \text{если } a_1 + b_1 = b_1; \\ b_1, & \text{если } a_1 + b_1 = a_1, \end{cases}$$

$$\text{т.е. } \mathbf{b} = \begin{cases} \mathbf{a}, & \text{если } \mathbf{a+b=b} \\ \mathbf{b}, & \text{если } \mathbf{a+b=a}; \end{cases}$$

7) S содержит элементы $\mathbf{1=1}$ и $\mathbf{0=0}$ такие, что для всякого элемента $\mathbf{a=a_1}$ из S:

$$a_1 + 0 = a_1, \text{ т.е. } \mathbf{a+0=a}$$

$$a_1 \cdot 1 = a_1, \text{ т.е. } \mathbf{a \cdot 1 = a}$$

$$a_1 \cdot 0 = 0, \text{ т.е. } \mathbf{a \cdot 0 = 0}$$

$$a_1 + 1 = 1, \text{ т.е. } \mathbf{a+1=1};$$

8) для каждого элемента $\mathbf{a=a_1}$ класс S содержит элемент $\overline{\mathbf{a}} = \overline{a_1}$ (дополнение элемента $\mathbf{a=a_1}$) такой, что:

$$a_1 + \overline{a_1} = 1, \text{ т.е. } \mathbf{a + \overline{a} = 1};$$

$$a_1 \cdot \overline{a_1} = 0, \text{ т.е. } \mathbf{a \cdot \overline{a} = 0}.$$

В каждой одномерной булевой алгебре имеют место:

9) законы поглощения:

$$a_1 \cdot (a_1 + b_1) = a_1, \text{ т.е. } \mathbf{a \cdot (a+b) = a},$$

$$a_1 + a_1 \cdot b_1 = a_1, \text{ т.е. } \mathbf{a + a \cdot b = a};$$

10) двойственность (законы де Моргана):

$$\overline{a_1 + b_1} = \overline{a_1} \cdot \overline{b_1}, \text{ т.е. } \overline{\mathbf{a + b}} = \overline{\mathbf{a}} \cdot \overline{\mathbf{b}},$$

$$\overline{a_1 \cdot b_1} = \overline{a_1} + \overline{b_1}, \text{ т.е. } \overline{\mathbf{a \cdot b}} = \overline{\mathbf{a}} + \overline{\mathbf{b}};$$

$$11) \bar{a}_1 = a_1, \text{ т.е. } \bar{\bar{a}} = a,$$

$$\bar{1} = 0, \text{ т.е. } \bar{\bar{1}} = 1,$$

$$\bar{0} = 1, \text{ т.е. } \bar{\bar{0}} = 0;$$

$$12) a_1 + \bar{a}_1 \cdot b_1 = a_1 + b_1, \text{ т.е. } a + \bar{a} \cdot b = a + b,$$

$$a_1 \cdot (\bar{a}_1 + b_1) = a_1 \cdot b_1, \text{ т.е. } a \cdot (\bar{a} + b) = a \cdot b;$$

$$13) a_1 \cdot b_1 + a_1 \cdot c_1 + b_1 \cdot \bar{c}_1 = a_1 \cdot c_1 + b_1 \cdot \bar{c}_1,$$

$$\text{т.е. } a \cdot b + a \cdot c + b \cdot \bar{c} = a \cdot c + b \cdot \bar{c},$$

$$(a_1 + b_1) \cdot (a_1 + c_1) \cdot (b_1 + \bar{c}_1) = (a_1 + c_1) \cdot (b_1 + \bar{c}_1),$$

$$\text{т.е. } (a + b) \cdot (a + c) \cdot (b + \bar{c}) = (a + c) \cdot (b + \bar{c}).$$

Двумерные Булевы алгебры

Двумерной Булевой алгеброй назовём класс S объектов $\mathbf{a}=(a_1, a_2)$, $\mathbf{b}=(b_1, b_2)$, $\mathbf{c}=(c_1, c_2)$, ..., в котором определены две бинарные операции, обозначаемые как (логические) сложение и умножение, со следующими свойствами:

для всех $\mathbf{a}=(a_1, a_2)$, $\mathbf{b}=(b_1, b_2)$, $\mathbf{c}=(c_1, c_2)$, ... из S имеют место;

1) (замкнутость) S содержит:

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2) = \mathbf{a} + \mathbf{b};$$

$$(a_1, a_2) \cdot (b_1, b_2) = (a_1 \cdot b_1, a_2 \cdot b_2) = \mathbf{a} \cdot \mathbf{b};$$

2) коммутативные законы:

$$(a_1, a_2) + (b_1, b_2) = (b_1, b_2) + (a_1, a_2);$$

$$(b_1, b_2) + (a_1, a_2) = (b_1 + a_1, b_2 + a_2), \text{ т.е. } \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a};$$

$$(a_1, a_2) \cdot (b_1, b_2) = (b_1, b_2) \cdot (a_1, a_2);$$

$$(b_1, b_2) \cdot (a_1, a_2) = (b_1 \cdot a_1, b_2 \cdot a_2), \text{ т.е. } \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a};$$

3) ассоциативные законы:

$$(a_1, a_2) + ((b_1, b_2) + (c_1, c_2)) = (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2));$$

$$((a_1, a_2) + (b_1, b_2)) + (c_1, c_2) = ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2);$$

$$\text{т.е. } \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c};$$

$$(a_1, a_2) \cdot ((b_1, b_2) \cdot (c_1, c_2)) = (a_1 \cdot (b_1 \cdot c_1), a_2 \cdot (b_2 \cdot c_2));$$

$$((a_1, a_2) \cdot (b_1, b_2)) \cdot (c_1, c_2) = ((a_1 \cdot b_1) \cdot c_1, (a_2 \cdot b_2) \cdot c_2),$$

$$\text{т.е. } \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c};$$

4) дистрибутивные законы:

$$(a_1, a_2) \cdot ((b_1, b_2) + (c_1, c_2)) = (a_1 \cdot (b_1 + c_1), a_2 \cdot (b_2 + c_2))$$

$$(a_1, a_2) \cdot (b_1, b_2) + (a_1, a_2) \cdot (c_1, c_2) = (a_1 \cdot b_1 + a_1 \cdot c_1, a_2 \cdot b_2 + a_2 \cdot c_2),$$

$$\text{т.е. } \mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) + (\mathbf{a} \cdot \mathbf{c}),$$

$$(a_1, a_2) + ((b_1, b_2) \cdot (c_1, c_2)) = (a_1 + (b_1 \cdot c_1), a_2 + (b_2 \cdot c_2))$$

$$((a_1, a_2) + (b_1, b_2)) \cdot (c_1, c_2) = ((a_1 + b_1) \cdot c_1, (a_2 + b_2) \cdot c_2),$$

$$\text{т.е. } \mathbf{a} + (\mathbf{b} \cdot \mathbf{c}) = (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c};$$

5) свойства идемпотентности:

$$(a_1, a_2) + (a_1, a_2) = (a_1, a_2), \text{ т.е. } \mathbf{a + a = a};$$

$$(a_1, a_2) \cdot (a_1, a_2) = (a_1, a_2), \text{ т.е. } \mathbf{a \cdot a = a};$$

6) свойства совместимости:

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2) = (a_1 + a_1 \cdot b_1, a_2 + a_2 \cdot b_2) = (a_1, a_2),$$

$$(a_1 \cdot b_1 + b_1, a_2 \cdot b_2 + b_2) = (b_1, b_2), \begin{cases} \text{если } (a_1, a_2) \cdot (b_1, b_2) = (b_1, b_2) \\ \text{если } (a_1, a_2) \cdot (b_1, b_2) = (a_1, a_2), \end{cases}$$

$$\text{т.е. } \mathbf{a + b = a}, \begin{cases} \text{если } \mathbf{a \cdot b = b}; \\ \mathbf{b}, \end{cases} \begin{cases} \text{если } \mathbf{a \cdot b = a}; \end{cases}$$

$$(a_1, a_2) \cdot (b_1, b_2) = (a_1 \cdot b_1, a_2 \cdot b_2) =$$

$$= (a_1 \cdot (a_1 + b_1), a_2 \cdot (a_2 + b_2)) = (a_1, a_2), \begin{cases} \text{если } (a_1, a_2) + (b_1, b_2) = (b_1, b_2), \\ ((a_1 + b_1) \cdot b_1, (a_2 + b_2) \cdot b_2) = (b_1, b_2), \end{cases} \begin{cases} \text{если } (a_1, a_2) + (b_1, b_2) = (a_1, a_2); \end{cases}$$

7) S содержит элементы $\mathbf{1} = (1, 1)$ и $\mathbf{0} = (0, 0)$ такие, что для всякого элемента: $\mathbf{a} = (a_1, a_2)$ из S:

$$(a_1, a_2) + (0, 0) = (a_1, a_2), \text{ т.е. } \mathbf{a + 0 = a},$$

$$(a_1, a_2) \cdot (1, 1) = (a_1, a_2), \text{ т.е. } \mathbf{a \cdot 1 = a},$$

$$(a_1, a_2) \cdot (0, 0) = (0, 0), \text{ т.е. } \mathbf{a \cdot 0 = 0},$$

$$(a_1, a_2) + (1, 1) = (1, 1), \text{ т.е. } \mathbf{a + 1 = 1};$$

8) для каждого элемента $\mathbf{a} = (a_1, a_2)$ класс S содержит элемент $\overline{\mathbf{a}} = (\overline{a_1}, \overline{a_2})$ (дополнение элемента $\mathbf{a} = (a_1, a_2)$) такой, что:

$$(a_1, a_2) + (\overline{a_1}, \overline{a_2}) = (a_1 + \overline{a_1}, a_2 + \overline{a_2}) = (1, 1), \text{ т.е. } \mathbf{a + \overline{a} = 1},$$

$$(a_1, a_2) \cdot (\overline{a_1}, \overline{a_2}) = (a_1 \cdot \overline{a_1}, a_2 \cdot \overline{a_2}) = (0, 0), \text{ т.е. } \mathbf{a \cdot \overline{a} = 0}.$$

В каждой двумерной булевой алгебре имеют место:

9) законы поглощения:

$$(a_1, a_2) \cdot ((a_1, a_2) + (b_1, b_2)) = (a_1 \cdot (a_1 + b_1), a_2 \cdot (a_2 + b_2)) = (a_1, a_2),$$

$$\text{т.е. } \mathbf{a \cdot (a + b) = a},$$

$$(a_1, a_2) + ((a_1, a_2) \cdot (b_1, b_2)) = (a_1 + (a_1 \cdot b_1), a_2 + (a_2 \cdot b_2)) = (a_1, a_2),$$

$$\text{т.е. } \mathbf{a + a \cdot b = a};$$

10) двойственность (законы де Моргана):

$$\overline{(a_1, a_2) + (b_1, b_2)} = \overline{(a_1 + b_1, a_2 + b_2)} = (\overline{a_1 + b_1}, \overline{a_2 + b_2}) = (\overline{a_1} \cdot \overline{b_1}, \overline{a_2} \cdot \overline{b_2}) = (\overline{a_1}, \overline{a_2}) \cdot (\overline{b_1}, \overline{b_2}),$$

$$\text{т.е. } \mathbf{\overline{a + b} = \overline{a} \cdot \overline{b}},$$

$$\overline{(a_1, a_2) \cdot (b_1, b_2)} = \overline{(a_1 \cdot b_1, a_2 \cdot b_2)} = (\overline{a_1 \cdot b_1}, \overline{a_2 \cdot b_2}) = (\overline{a_1} + \overline{b_1}, \overline{a_2} + \overline{b_2}) = (\overline{a_1}, \overline{a_2}) + (\overline{b_1}, \overline{b_2}),$$

$$\text{т.е. } \mathbf{\overline{a \cdot b} = \overline{a} + \overline{b}};$$

$$11) \quad \overline{\overline{(a_1, a_2)}} = \overline{(\overline{a_1}, \overline{a_2})} = (a_1, a_2), \text{ т.е. } \mathbf{\overline{\overline{a}} = a},$$

$$\overline{\overline{(1, 1)}} = \overline{(\overline{1}, \overline{1})} = (0, 0), \text{ т.е. } \mathbf{\overline{\overline{1}} = 0},$$

$$\overline{\overline{(0, 0)}} = \overline{(\overline{0}, \overline{0})} = (1, 1), \text{ т.е. } \mathbf{\overline{\overline{0}} = 1};$$

$$\begin{aligned}
12) \quad & (a_1, a_2) + (\bar{a}_1, \bar{a}_2) \cdot (b_1, b_2) = (a_1 + \bar{a}_1 \cdot b_1, a_2 + \bar{a}_2 \cdot b_2) = \\
& = (a_1 + b_1, a_2 + b_2) = (a_1, a_2) + (b_1, b_2), \text{ т.е. } \mathbf{a} + \mathbf{a} \cdot \mathbf{b} = \mathbf{a} + \mathbf{b}, \\
& (a_1, a_2) \cdot ((\bar{a}_1, \bar{a}_2) + (b_1, b_2)) = (a_1 \cdot (\bar{a}_1 + b_1), a_2 \cdot (\bar{a}_2 + b_2)) = \\
& = (a_1 \cdot b_1, a_2 \cdot b_2) = (a_1, a_2) \cdot (b_1, b_2), \text{ т.е. } \mathbf{a} \cdot (\mathbf{a} + \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}; \\
13) \quad & (a_1, a_2) \cdot (b_1, b_2) + (a_1, a_2) \cdot (c_1, c_2) + \\
& + (b_1, b_2) \cdot (c_1, c_2) = (a_1 \cdot b_1 + a_1 \cdot c_1 + b_1 \cdot c_1, a_2 \cdot b_2 + a_2 \cdot c_2 + b_2 \cdot c_2) = \\
& = (a_1 \cdot c_1 + b_1 \cdot c_1, a_2 \cdot c_2 + b_2 \cdot c_2), \\
& \text{т.е. } \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}; \\
& ((a_1, a_2) + (b_1, b_2)) \cdot ((a_1, a_2) + (c_1, c_2)) \cdot ((b_1, b_2) + (c_1, c_2)) = \\
& = ((a_1 + b_1) \cdot (a_1 + c_1) \cdot (b_1 + c_1), \\
& (a_2 + b_2) \cdot (a_2 + c_2) \cdot (b_2 + c_2)) = ((a_1 + c_1) \cdot (b_1 + c_1), (a_2 + c_2) \cdot (b_2 + c_2)), \\
& \text{т.е. } (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{c}) \cdot (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{c}) \cdot (\mathbf{b} + \mathbf{c}).
\end{aligned}$$

п-мерные Булевы алгебры

п-мерной Булевой алгеброй назовём класс S объектов $\mathbf{a} = (a_1, a_2, \dots, a_n)$, $\mathbf{b} = (b_1, b_2, \dots, b_n)$, $\mathbf{c} = (c_1, c_2, \dots, c_n)$, ..., в котором определены две бинарные операции, обозначаемые как (логические) сложение и умножение, со следующими свойствами:

для всех $\mathbf{a} = (a_1, a_2, \dots, a_n)$, $\mathbf{b} = (b_1, b_2, \dots, b_n)$, $\mathbf{c} = (c_1, c_2, \dots, c_n)$, ... из S имеют место:

1) (замкнутость) S содержит:

$$\begin{aligned}
& (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = \mathbf{a} + \mathbf{b}; \\
& (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n) = \mathbf{a} \cdot \mathbf{b};
\end{aligned}$$

2) коммутативные законы:

$$\begin{aligned}
& (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \\
& (b_1, b_2, \dots, b_n) + (a_1, a_2, \dots, a_n) = (b_1 + a_1, b_2 + a_2, \dots, b_n + a_n), \\
& \text{т.е. } \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \\
& (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n) \\
& (b_1, b_2, \dots, b_n) \cdot (a_1, a_2, \dots, a_n) = (b_1 \cdot a_1, b_2 \cdot a_2, \dots, b_n \cdot a_n), \\
& \text{т.е. } \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a};
\end{aligned}$$

3) ассоциативные законы:

$$\begin{aligned}
& (a_1, a_2, \dots, a_n) + ((b_1, b_2, \dots, b_n) + (c_1, c_2, \dots, c_n)) = \\
& = (a_1 + (b_1 + c_1), a_2 + (b_2 + c_2), \dots, a_n + (b_n + c_n)) \\
& ((a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)) + (c_1, c_2, \dots, c_n) = \\
& = ((a_1 + b_1) + c_1, (a_2 + b_2) + c_2, \dots, (a_n + b_n) + c_n), \\
& \text{т.е. } \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}; \\
& (a_1, a_2, \dots, a_n) \cdot ((b_1, b_2, \dots, b_n) \cdot (c_1, c_2, \dots, c_n)) = \\
& = (a_1 \cdot (b_1 \cdot c_1), a_2 \cdot (b_2 \cdot c_2), \dots, a_n \cdot (b_n \cdot c_n)) \\
& ((a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n)) \cdot (c_1, c_2, \dots, c_n) = \\
& = ((a_1 \cdot b_1) \cdot c_1, (a_2 \cdot b_2) \cdot c_2, \dots, (a_n \cdot b_n) \cdot c_n), \\
& \text{т.е. } \mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c};
\end{aligned}$$

4) дистрибутивные законы:

$$\begin{aligned}
 &(a_1, a_2, \dots, a_n) \cdot ((b_1, b_2, \dots, b_n) + (c_1, c_2, \dots, c_n)) = \\
 &= (a_1 \cdot (b_1 + c_1), a_2 \cdot (b_2 + c_2), \dots, a_n \cdot (b_n + c_n)) \\
 &(a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) + (a_1, a_2, \dots, a_n) \cdot (c_1, c_2, \dots, c_n) = \\
 &= (a_1 \cdot b_1 + a_1 \cdot c_1, a_2 \cdot b_2 + a_2 \cdot c_2, \dots, a_n \cdot b_n + a_n \cdot c_n), \\
 &\text{т.е. } \mathbf{a \cdot (b + c) = (a \cdot b) + (a \cdot c)}, \\
 &(a_1, a_2, \dots, a_n) + ((b_1, b_2, \dots, b_n) \cdot (c_1, c_2, \dots, c_n)) = \\
 &= (a_1 + (b_1 \cdot c_1), a_2 + (b_2 \cdot c_2), \dots, a_n + (b_n \cdot c_n)) \\
 &((a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)) \cdot ((a_1, a_2, \dots, a_n) + (c_1, c_2, \dots, c_n)) = \\
 &= ((a_1 + b_1) \cdot (a_1 + c_1), (a_2 + b_2) \cdot (a_2 + c_2), \dots, (a_n + b_n) \cdot (a_n + c_n)), \\
 &\text{т.е. } \mathbf{a + (b \cdot c) = (a + b) \cdot (a + c)};
 \end{aligned}$$

5) свойства идемпотентности:

$$\begin{aligned}
 &(a_1, a_2, \dots, a_n) + (a_1, a_2, \dots, a_n) = (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a + a = a} \\
 &(a_1, a_2, \dots, a_n) \cdot (a_1, a_2, \dots, a_n) = (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a \cdot a = a};
 \end{aligned}$$

6) свойства совместимости:

$$\begin{aligned}
 &(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = \\
 &\left\{ \begin{aligned} &(a_1 + a_1 \cdot b_1, a_2 + a_2 \cdot b_2, \dots, a_n + a_n \cdot b_n) = (a_1, a_2, \dots, a_n), \\ &= \text{если } (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (b_1, b_2, \dots, b_n) \\ &(a_1 \cdot b_1 + b_1, a_2 \cdot b_2 + b_2, \dots, a_n \cdot b_n + b_n) = (b_1, b_2, \dots, b_n), \\ &\text{если } (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (a_1, a_2, \dots, a_n), \end{aligned} \right. \\
 &\text{т.е. } \mathbf{a + b = \begin{cases} a, & \text{если } a \cdot b = b \\ b, & \text{если } a \cdot b = a \end{cases}} \\
 &(a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) = (a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n) = \\
 &\left\{ \begin{aligned} &(a_1 \cdot (a_1 + b_1), a_2 \cdot (a_2 + b_2), \dots, a_n \cdot (a_n + b_n)) = (a_1, a_2, \dots, a_n), \\ &= \text{если } (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (b_1, b_2, \dots, b_n) \\ &((a_1 + b_1) \cdot b_1, (a_2 + b_2) \cdot b_2, \dots, (a_n + b_n) \cdot b_n) = (b_1, b_2, \dots, b_n), \\ &\text{если } (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1, a_2, \dots, a_n), \end{aligned} \right. \\
 &\text{т.е. } \mathbf{a \cdot b = \begin{cases} a, & \text{если } a + b = b \\ b, & \text{если } a + b = a; \end{cases}}
 \end{aligned}$$

7) S содержит элементы $\mathbf{1} = (1, 1, \dots, 1)$ и $\mathbf{0} = (0, 0, \dots, 0)$ такие, что для всякого элемента $\mathbf{a} = (a_1, a_2, \dots, a_n)$ из S:

$$\begin{aligned}
 &(a_1, a_2, \dots, a_n) + (0, 0, \dots, 0) = (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a + 0 = a}; \\
 &(a_1, a_2, \dots, a_n) \cdot (1, 1, \dots, 1) = (a_1, a_2, \dots, a_n), \text{ т.е. } \mathbf{a \cdot 1 = a}; \\
 &(a_1, a_2, \dots, a_n) \cdot (0, 0, \dots, 0) = (0, 0, \dots, 0), \text{ т.е. } \mathbf{a \cdot 0 = 0}; \\
 &(a_1, a_2, \dots, a_n) + (1, 1, \dots, 1) = (1, 1, \dots, 1), \text{ т.е. } \mathbf{a + 1 = 1};
 \end{aligned}$$

8) для каждого элемента $\mathbf{a} = (a_1, a_2, \dots, a_n)$ класс S содержит элемент $\mathbf{\bar{a} = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n)}$ (дополнение элемента $\mathbf{a} = (a_1, a_2, \dots, a_n)$) такой, что

$$\begin{aligned}
 &(a_1, a_2, \dots, a_n) + (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) = \\
 &= (a_1 + \bar{a}_1, a_2 + \bar{a}_2, \dots, a_n + \bar{a}_n) = (1, 1, \dots, 1), \\
 &\text{т.е. } \mathbf{a + \bar{a} = 1}; \\
 &(a_1, a_2, \dots, a_n) \cdot (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_n) = \\
 &= (a_1 \cdot \bar{a}_1, a_2 \cdot \bar{a}_2, \dots, a_n \cdot \bar{a}_n) = (0, 0, \dots, 0), \\
 &\text{т.е. } \mathbf{a \cdot \bar{a} = 0}.
 \end{aligned}$$

В каждой n -мерной Булевой алгебре имеют место:

9) законы поглощения:

$$(a_1, a_2, \dots, a_n) \cdot ((a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n)) = (a_1, a_2, \dots, a_n) \cdot (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = (a_1, a_2, \dots, a_n),$$

т.е. **$a \cdot (a + b) = a$** ;

$$(a_1, a_2, \dots, a_n) + ((a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n)) = (a_1 + (a_1 \cdot b_1), a_2 + (a_2 \cdot b_2), \dots, a_n + (a_n \cdot b_n)) = (a_1, a_2, \dots, a_n),$$

т.е. **$a + a \cdot b = a$** ;

10) двойственность (законы де Моргана):

$$(\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) + (\overline{b_1}, \overline{b_2}, \dots, \overline{b_n}) = \overline{(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)} = \overline{(a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n)} = \overline{(a_1, a_2, \dots, a_n)} \cdot \overline{(b_1, b_2, \dots, b_n)},$$

т.е. **$a + b = \overline{a \cdot b}$** ;

$$(\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) \cdot (\overline{b_1}, \overline{b_2}, \dots, \overline{b_n}) = \overline{(a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n)} = \overline{(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)} = \overline{(a_1, a_2, \dots, a_n)} + \overline{(b_1, b_2, \dots, b_n)},$$

т.е. **$a \cdot b = \overline{a + b}$** ;

$$11) \quad (\overline{\overline{a_1}}, \overline{\overline{a_2}}, \dots, \overline{\overline{a_n}}) = (\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) = (a_1, a_2, \dots, a_n), \text{ т.е. } \overline{\overline{a}} = a,$$

$$(\overline{\overline{1}}, \overline{\overline{1}}, \dots, \overline{\overline{1}}) = (\overline{1}, \overline{1}, \dots, \overline{1}) = (0, 0, \dots, 0), \text{ т.е. } \overline{\overline{1}} = 0,$$

$$(\overline{\overline{0}}, \overline{\overline{0}}, \dots, \overline{\overline{0}}) = (\overline{0}, \overline{0}, \dots, \overline{0}) = (1, 1, \dots, 1), \text{ т.е. } \overline{\overline{0}} = 1;$$

$$12) \quad (a_1, a_2, \dots, a_n) + (\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) \cdot (b_1, b_2, \dots, b_n) = (a_1 + a_1 \cdot b_1, a_2 + a_2 \cdot b_2, \dots, a_n + a_n \cdot b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) = (a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n), \text{ т.е. } a + a \cdot b = a + b;$$

$$(a_1, a_2, \dots, a_n) \cdot ((\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) + (b_1, b_2, \dots, b_n)) = (a_1 \cdot (a_1 + b_1), a_2 \cdot (a_2 + b_2), \dots, a_n \cdot (a_n + b_n)) = (a_1 \cdot b_1, a_2 \cdot b_2, \dots, a_n \cdot b_n) = (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n), \text{ т.е. } a \cdot (a + b) = a \cdot b;$$

$$13) \quad (a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_n) + (\overline{a_1}, \overline{a_2}, \dots, \overline{a_n}) \cdot (c_1, c_2, \dots, c_n) + (b_1, b_2, \dots, b_n) \cdot (\overline{c_1}, \overline{c_2}, \dots, \overline{c_n}) = (a_1 \cdot b_1 + a_1 \cdot c_1 + b_1 \cdot \overline{c_1}, a_2 \cdot b_2 + a_2 \cdot c_2 + b_2 \cdot \overline{c_2}, \dots, a_n \cdot b_n + a_n \cdot c_n + b_n \cdot \overline{c_n}) = (\overline{a_1 \cdot c_1 + b_1 \cdot \overline{c_1}}, \overline{a_2 \cdot c_2 + b_2 \cdot \overline{c_2}}, \dots, \overline{a_n \cdot c_n + b_n \cdot \overline{c_n}}) = (\overline{a_1} \cdot \overline{c_1} + \overline{b_1} \cdot c_1, \overline{a_2} \cdot \overline{c_2} + \overline{b_2} \cdot c_2, \dots, \overline{a_n} \cdot \overline{c_n} + \overline{b_n} \cdot c_n) = ((\overline{a_1} + \overline{b_1}) \cdot (\overline{a_1} + \overline{c_1}), (\overline{a_2} + \overline{b_2}) \cdot (\overline{a_2} + \overline{c_2}), \dots, (\overline{a_n} + \overline{b_n}) \cdot (\overline{a_n} + \overline{c_n})) = ((a_1 + b_1) \cdot (a_1 + c_1), (a_2 + b_2) \cdot (a_2 + c_2), \dots, (a_n + b_n) \cdot (a_n + c_n)),$$

т.е. **$(a + b) \cdot (a + c) \cdot (b + c) = (a + c) \cdot (b + c)$** .

Таким образом, свойства многомерных Булевых алгебр повторяют свойства одномерной Булевой алгебры.

Список литературы

1. Корн Г., Корн Т. Справочник по математике (для научных работников и инженеров). – М.: Наука, 1977. – 832 с.

РАЗДЕЛ 2

ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

УДК 004.896:004.021

В.В. Бегляров

СИСТЕМАТИЧЕСКИЙ АНАЛИЗ ПРИНЦИПОВ ЭВОЛЮЦИОННОГО МОДЕЛИРОВАНИЯ

В статье отражено настоящее состояние дел в области генетических алгоритмов, приведены дальнейшие пути развития этого направления. В работе описывается метод поиска оптимальных решений, основанный на генетическом подходе, приведено описание как традиционных моделей эволюции, так и разработанных автором. Отражены основные понятия генетического алгоритма, приведён перечень и описание основных генетических операторов. Перечислены основные проблемы применения генетических алгоритмов, а также возможные пути их решения.

Введение. В настоящее время перед наукой всё больше ставятся задачи, математическое описание которых либо отсутствует вовсе, либо не отражает проблему в полной мере. Такие задачи невозможно или практически невозможно решать с помощью традиционных методов поиска решений. Поэтому в настоящее время существует потребность в нахождении и использовании нетрадиционных, но более эффективных методов поиска решений.

К задачам, требующим использования новых нетривиальных методов, можно отнести: задачи многокритериальной оптимизации, задачи принятия решений, комбинаторные задачи на графах, задачи адаптации (обучения нейронных сетей, адаптация агентов), задачи подбора параметров, задачи составления расписаний и т.д. Решение таких задач очень важно для науки, т.к. нахождение более эффективных методов решения способно вывести науку и технику на новый уровень. Например, решение задач поиска параметров очень важно для современных САПР, одной из таких решаемых задач является подбор параметров элементов электрической цепи в процедурах параметрической оптимизации.

На сегодняшний день в науке и технике для решения обширного круга задач применяются следующие методы, относящиеся к искусственному интеллекту: нейронные сети, генетические алгоритмы, эволюционные алгоритмы, нечёткие множества и логика. Вышеперечисленные методологии объединяются под зонтичным термином «мягкие вычисления», введённым в 1994 г. Л. Заде. В данной статье речь пойдёт о генетических алгоритмах (ГА). Свою популярность они завоевали простотой реализации и достаточным уровнем абстракции естественных процессов эволюции [1].

Первый ГА был описан Д. Гольдбергом на основе работ Д. Холланда. Основу ГА составляет направленный поиск, принцип работы которого базируется на идеях эволюции живой природы.

Растущий объём публикаций по этой тематике свидетельствует о большом интересе к использованию генетических алгоритмов. Результаты в данной области были достигнуты следующими научными школами: Д. Гольдберга, Д. Холланда, Д. Козы, Д.И. Батищева, В.М. Курейчика, Д.А. Поспелова, Н.П. Норенкова и др.

1. Формальное описание генетического алгоритма и основные модели эволюции

Основу ГА составляют три основные операции: скрещивание, мутация, селекция.

Генетический алгоритм можно описать следующим образом:

$$ГА = \langle P^0, \lambda, l, s, p, f, t \rangle,$$

где $P^0 = (a^0_1, \dots, a^0_\lambda)$ – исходная популяция, a^0_i – решение задачи, представленное в виде хромосомы; λ – целое число (размер популяции); l – целое число (длина каждой хромосомы популяции); s – оператор отбора; p – отображение, определяющее рекомбинацию (кроссинговер, мутация); f – функция оптимальности; t – критерий останова.

В настоящее время генетические алгоритмы являются наиболее популярными эволюционными алгоритмами, это связано с простотой реализации и достаточным уровнем абстракции естественных процессов эволюции, который реализован в ГА [2–4].

Наиболее известными являются следующие модели эволюции: Дарвина, Ламарка, де Фризе, Поппера, Дубинина, Кимура. Приведём описание наиболее часто используемых моделей.

Модель эволюции Ч. Дарвина. Условная упрощённая схема модели эволюции Ч. Дарвина приведена на рисунке 1.

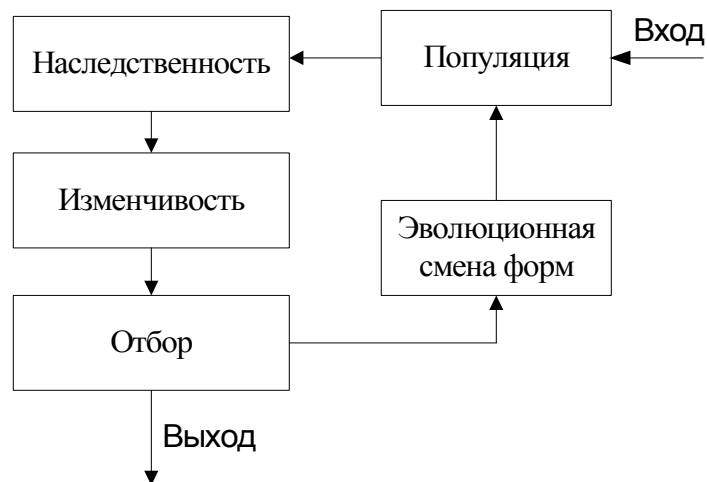


Рис. 1. Условная упрощённая схема модели эволюции Ч. Дарвина

Теория Чарльза Дарвина (дарвинизм), известная под названием теории естественного отбора, является наиболее законченной теорией эволюции. Согласно этой теории основными условиями эволюции являются:

- наследственная изменчивость, т.е. мутирование, понимаемое как предпосылка эволюции, её материал;
- борьба за существование как контролирующий и направляющий фактор эволюции;
- естественный отбор как преобразующий фактор эволюции.

Основными факторами естественного отбора являются механизм отбора; режим размножения; процессы индивидуального развития. Мутации являются теми наследственными изменениями, которые либо отдельно, либо совместно определяют изменения свойств, признаков, особенностей или норм реакции организмов. В центре внимания эволюционной теории находятся не отдельные организмы, а биологические виды и внутривидовые группировки (популяции).

Модель эволюции Ч. Дарвина – это условная структура, реализующая процесс, посредством которого особи некоторой популяции, имеющие более высокое функциональное значение, получают большую возможность для воспроизведения потомков, чем «слабые» особи. Такой механизм часто называют методом «выживания сильнейших» [3–7].

ГА, использующий данную модель, способен найти глобальный оптимум, игнорируя локальные оптимумы и разрывы целевой функции. К недостаткам можно отнести достаточную вероятность быстрой сходимости популяции, что может привести к снижению эффективности.

Модель эволюции Ж. Ламарка. На рисунке 2 представлена условная упрощённая схема модели эволюции Ж. Ламарка.

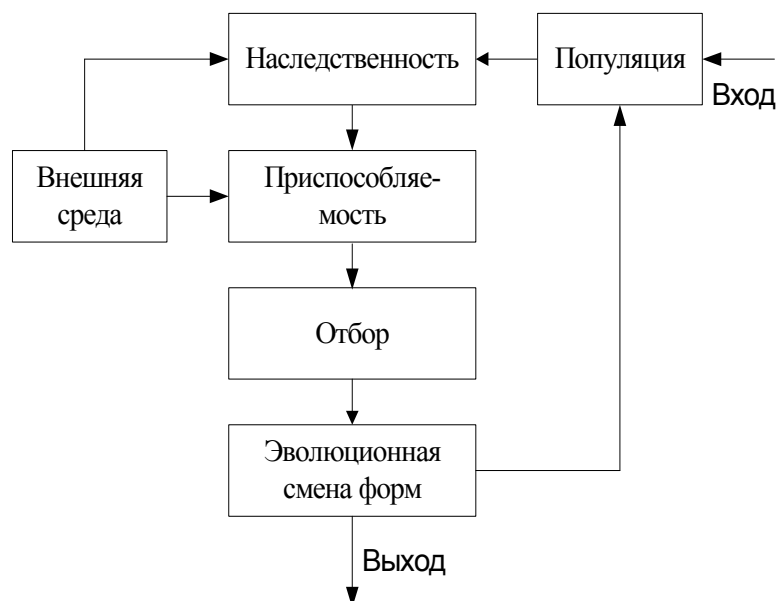


Рис. 2. Условная упрощённая схема модели эволюции Ж. Ламарка

Согласно модели Ламарка основным фактором эволюции является наследование приобретённых признаков, а материалом эволюции является приспособление к внешней среде. Ламарк объясняет одну из особенностей эволюции органического мира приспособляемостью. Прогрессивную эволюцию, появление форм, более сложных и совершенных, он объяснял «законом градаций», стремлением живых существ усложнять свою структуру. Согласно Ж. Ламарку, виды эволюционируют, приспособляясь и усложняясь, потому что у них существуют свойства приспособляемости и усложняемости.

В терминах ГА нового индивида предлагается создавать путём рекомбинации генов. Основным же отличием эволюции, по Ламарку, является приспособление к внешней среде. В ГА реализуется это как улучшение фенотипа (а следовательно, и генотипа) индивида. Для этого возможно использование локальных или других методов поиска. Однако такое обучение приводит к разрушению процесса схемообразования.

Локализованный поиск Ламарка более подходит для исследования локальных областей популяции, в отличие от глобального перебора ГА. Эволюционная модель Ламарка оказывается наиболее эффективной, когда популяция имеет сходимость в область локального минимума в отличие от стандартного ГА.

Для уменьшения негативного эффекта от применения эволюции Ламарка, с точки зрения разрушения наследственных схем, была предложена методика, основанная на биологическом «эффекте Болдуина». Её суть состоит в улучшении фенотипа хромосомы и запоминании полученного значения без изменения генотипа хромосомы. ГА с «эффектом Болдуина» уменьшают преждевременную сходимость по сравнению с ГА Ламарка, незначительно уменьшая скорость нахождения решения [3, 5–7].

ГА, использующий модель Ламарка, способен найти глобальный оптимум, минуя разрывы целевой функции. К достоинствам можно отнести адаптацию к изменяющимся условиям, к недостаткам – немалую возможность сходимости к локальному оптимуму, а также преждевременную сходимость популяции.

Модель эволюции Г. де Фриза. В основе этой модели лежит моделирование социальных и географических катастроф, приводящих к резкому изменению видов и популяций. Эволюция, таким образом, представляет собой последовательность скачков в развитии популяции без предварительного накопления количественных изменений в эволюционных процессах. Он проявляется один раз ориентировочно в несколько тысяч поколений. Основная идея его состоит во внесении глобальных изменений в генотип на момент катастрофы [3, 5–7].

На рисунке 3 представлена условная упрощённая схема модели эволюции де Фриза.

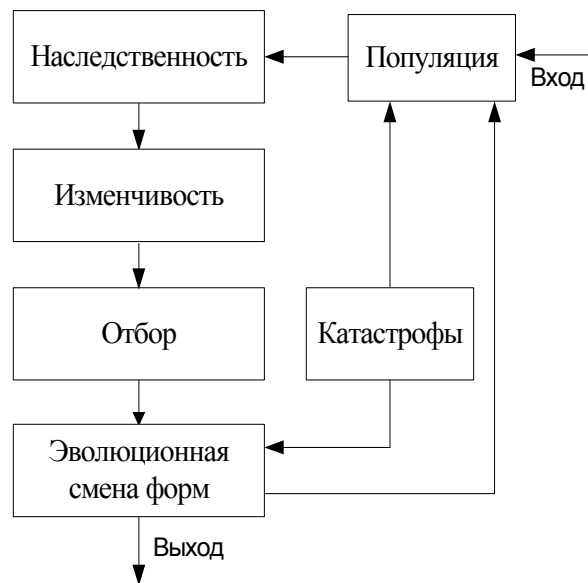


Рис. 3. Условная упрощённая схема модели эволюции де Фриза

ГА, использующий модель де Фриза, способен найти глобальный оптимум, игнорируя разрывы целевой функции. К достоинствам можно отнести механизм сильного взбалтывания популяции. Последнее преимущество нередко может являться и большим недостатком, так как способно отвести популяцию от нужного вектора развития.

Модель К. Поппера – это условная структура, реализующая иерархическую систему гибких механизмов управления, в которых мутация интерпретируется как метод случайных проб и ошибок, а отбор – как один из способов управления с помощью устранения ошибок при взаимодействии с внешней средой. Эволюция К. Поппера излагается в виде двенадцати тезисов. Основные из них:

- проблемы эволюции всегда решаются методом проб и ошибок;
- устранение ошибок может осуществляться либо путём полного устранения неудачных форм, либо в виде эволюции механизмов управления;
- популяция использует тот механизм управления, который выработался в процессе эволюции;
- популяция является пробным решением, анализируемым в процессе эволюции, выбирающим окружающую среду и преобразующим её;
- эволюционная последовательность событий представляется в виде последовательности $F_1 \rightarrow TS \rightarrow EE \rightarrow F_2$, где F_1 – исходная проблема, TS – пробные решения, EE – устранение ошибок, F_2 – новая проблема.

В отличие от эволюции Ч. Дарвина, где существует одна проблема – «выживание сильнейших», – в эволюции К. Поппера существуют и другие проблемы: воспроизводство, избавление от лишнего потомства и т.п. Согласно К. Попперу, естественные системы исследуют окружающую среду и активно получают из неё информацию. Процесс выбора лучшей индивидуальности в данной эволюции может являться процессом отбора (селекции), а отбор из некоторого множества случайных событий не обязан быть случайным. На рисунке 4 приведена условная упрощённая схема модели эволюции К. Поппера [3, 5–7].

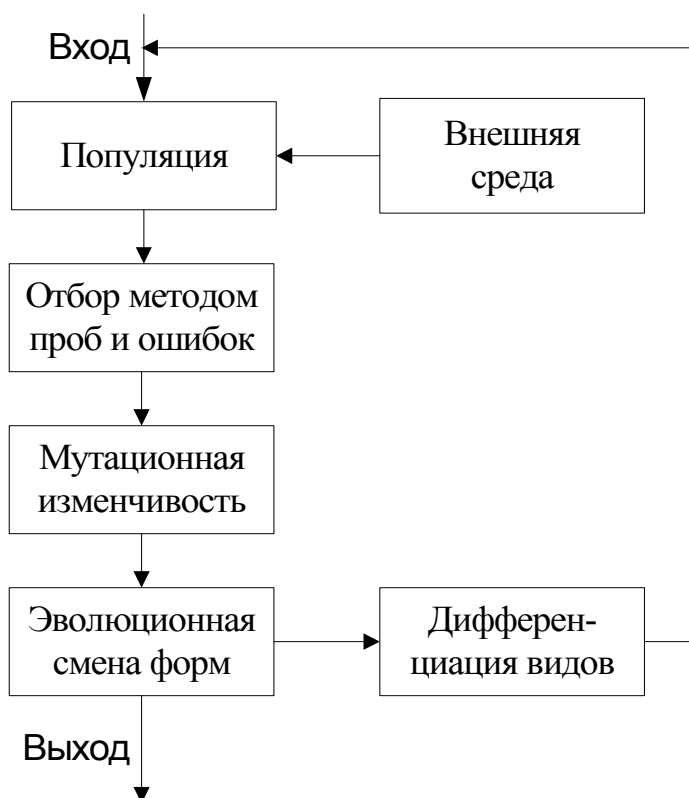


Рис. 4. Условная упрощённая схема модели эволюции К. Поппера

ГА, использующий данную модель, способен найти глобальный оптимум, минуя разрывы целевой функции. Также к достоинствам можно отнести адаптацию к изменяющимся условиям, а к недостаткам – сложность реализации, а также то, что в большей мере используется случайный поиск и процесс сходимости может быть достаточно долгим.

В таблице приведено сравнение моделей эволюции на основе экспертных оценок.

Сводная таблица моделей эволюции

Модель	Достижение глобального оптимума	Игнорирование разрывов целевой функции	Взбалтывание популяции (5 баллов)	Скорость сходимости (5 баллов)
Ч. Дарвина	+	+	3	4
Ламарка	+	+	3	4
Де Фриза	+	+	5	3
К. Поппера	+	+	3	2
Кимуры	+	+	2	4
Дубинина	+	+	4	3

Для повышения эффективности работы генетических алгоритмов необходимо использовать сочетания разных моделей эволюции. Как известно, применение методов, соответствующих одной научной парадигме, для решения сложных проблем далеко не всегда является эффективным. Гораздо лучший результат дают гибридные архитектуры, т.к. объединение нескольких парадигм позволяет компенсировать их слабые стороны. Комбинируя различные подходы, можно обойти недостатки, присущие каждому из них в отдельности. Поэтому одной из ведущих тенденций стало распространение интегрированных и гибридных систем. Подобные системы состоят из различных элементов (компонентов), объединённых в интересах достижения поставленных целей. Интеграция и гибридизация различных методов и информационных технологий позволяет решать сложные задачи, которые невозможно решить на основе каких-либо отдельных методов или технологий. Одним из примеров таких гибридных систем являются системы, сочетающие в себе нейронные системы, нечёткую логику и генетические алгоритмы. Такие системы хорошо решают задачи, связанные с обработкой неполных, неточных и недостоверных данных [2, 8–10].

Наибольшую сложность в построении таких систем представляют алгоритмы переключения между однородными парадигмами, а также создание эффективных алгоритмов, способных сочетать лучшие качества каждой из неоднородных парадигм. Также следует отметить, что в настоящее время используется крайне узкий класс эволюционных теорий, причём объединяются только отдельные элементы различных теорий эволюции. Представляется необходимым расширение списка эволюционных учений, применяемых в ГА.

Одним из наиболее популярных методов повышения эффективности работы генетических алгоритмов является использование многопопуляционных алгоритмов, в которых каждая популяция может развиваться согласно своей модели эволюции и иметь свой, отличный от других, набор операторов и параметров. На рисунке 5 представлена разработанная автором модель многопопуляционного алгоритма.

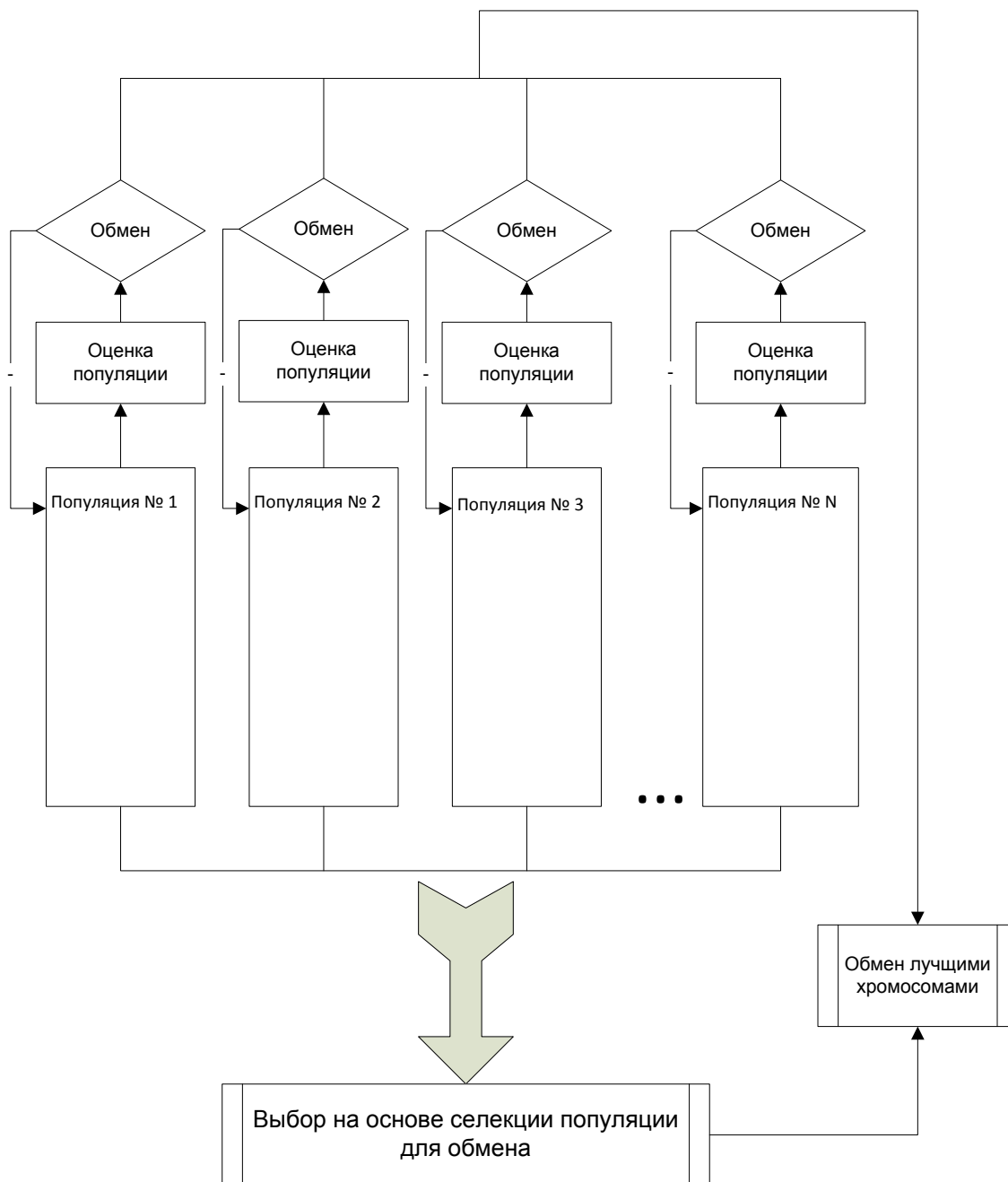


Рис. 5. Модель многопопуляционного алгоритма

Сущность предложенной модели заключается в одновременном развитии нескольких популяций, каждая из которых развивается по собственным законам. Для оценки качества популяции использовались следующие показатели:

– разнообразие генотипа (GD):

а) хэммингово расстояние (при битовом кодировании);

б) евклидово расстояние (при вещественном кодировании).

$$GD = \frac{\bar{d} - d_{\min}}{d_{\max} - d_{\min}},$$

где $\bar{d} = \frac{1}{N} \sum_{i=1}^N d(C_{best}, C_i)$, $d_{\max} = \max\{d(C_{best}, C_i) | C_i \in P\}$;

$d_{\min} = \min\{d(C_{best}, C_i) | C_i \in P\}$,

где P – популяция хромосом; C_{best} – лучшая хромосома. $GD \in [0, 1]$. Если величина GD мала, то большинство хромосом сконцентрировано вокруг лучшей хромосомы и сходимость достигнута.

в) дисперсия значений генов. Предлагается различать дисперсию между генами (BC) и дисперсию между аллелями (BA).

Для двоичного кодирования, если рассматривать популяцию как матрицу, в которой строки – хромосомы, а столбцы – аллели, получим:

$$BC = \frac{\frac{\sum_i s_i^2}{L} - \frac{s^2}{L * N}}{N - 1}, BA = \frac{\frac{\sum_j s_j^2}{N} - \frac{s^2}{L * N}}{L - 1},$$

где N – размер популяции; L – длина хромосомы; S – сумма генов «1»; S_i – сумма над рядом i ; S_j – сумма правее столбца j . BA и BC обладают следующими свойствами:

- 1) если популяция гомогенна 0 или 1, то BC и BA равны 0;
- 2) когда все хромосомы в популяции идентичны, то $BC = 0$, а BA содержит константу, определяющую количество 1 в хромосоме;
- 3) BA не зависит от оператора кроссинговера;
- 4) BC и BA всегда меньше, чем $\frac{L}{N}$ и $\frac{N}{L}$ соответственно.

Для вещественного кодирования:

$$VAC = \frac{\sum_{i=1}^N (\bar{S}_i - \bar{S})^2}{N} \text{ – разнообразие генов;}$$

$$AVA = \frac{\sum_{j=1}^N \sum_{i=1}^L (S_{ij} - \bar{S}_i)^2}{LN} \text{ – разнообразие аллелей;}$$

$S_i, i = 1, 2, \dots, N$ – значение i -й хромосомы;

$S_j, j = 1, 2, \dots, L$ – вектор j -х генов популяции;

S_{ij} – j -й ген в i -й хромосоме;

$$\bar{S}_i = \frac{\sum_{j=1}^L S_{ij}}{L}, \quad \bar{S} = \frac{\sum_{j=1}^N \sum_{i=1}^L S_{ij}}{LN}, \quad \bar{S}_j = \frac{\sum_{i=1}^N S_{ij}}{N}.$$

Меры VAC и AVA не зависят от обмена генами двух хромосом в популяции, и когда все хромосомы в популяции почти одинаковы, AVA и VAC имеют малые значения.

– разнообразие фенотипа. $PD = \frac{f_{best}}{\bar{f}}$, где $\bar{f}, f_{best}$ – среднее и лучшее

значение функции пригодности. Величина $PD \in [0, 1]$. Если значение PD около 1, то сходимость достигнута, а если PD около 0, то в текущий момент можно говорить о высоком уровне разнообразия в популяции.

– изменение лучшей хромосомы:

$$\Delta f_{bestk} = \frac{f_{besti} - f_{besti-k}}{\Delta Y},$$

где $f_{besti}, f_{besti-k}$ – лучшее значение функции пригодности на i -м и $(i - k)$ -м поколениях; ΔY – точность нахождения решения; k – количество анализируемых поколений.

Если Δf_{best} мало, то прогресс популяции отсутствует.

– среднее изменение функции пригодности популяции:

$$\Delta f_{avk} = \frac{f_{avi} - f_{avi-k}}{\Delta Y},$$

где f_{avi}, f_{avi-k} – среднее значение функции пригодности на i -м и $(i - k)$ -м поколениях; ΔY – точность нахождения решения; k – количество анализируемых поколений. Если Δf_{avk} мало, то популяции не развиваются [8].

Анализируя вышеперечисленные параметры, принимаем решение, необходим ли данной популяции обмен хромосомами с другой популяцией. Выбор популяции для обмена в соответствии с выбранной моделью селекции происходит из оставшихся популяций. После миграции хромосом процесс развития популяций продолжается.

Данная модель является более эффективной, чем однопопуляционная, т.к. при одновременном развитии многих популяций, каждая из которых развивается по собственной модели, происходит совмещение достоинств всех моделей и подходов и одновременно уменьшается влияние их недостатков. Ещё одним достоинством такого подхода являются исследования гораздо большего пространства решений.

2. Обобщённая структура генетического алгоритма и генетические операторы

Обобщённую структуру генетического алгоритма можно описать следующей последовательностью шагов:

Шаг 1. Генерация начальной популяции согласно выбранным критериям.

Шаг 2. Отбор хромосом согласно выбранному виду селекции для проведения генетических операторов.

Шаг 3. Выполнение выбранных генетических операторов.

Шаг 4. Приведение популяции к исходному размеру согласно выбранным параметрам.

Шаг 5. Сбор статистических параметров развития популяции.

Шаг 6. Проверка выбранных условий завершения процесса. В случае их выполнения переходим к шагу 7, иначе – к шагу 2.

Шаг 7. Конец.

Эффективность работы любого генетического алгоритма неразрывно связана с правильным подбором набора генетических операторов и параметров работы как самого алгоритма в целом, так и каждого оператора в отдельности. Ниже будет приведён список генетических операторов с их описанием.

Оператор селекции

Ниже приводится описание использованных операторов селекции:

- селекция на основе рулетки – это самый простой и наиболее используемый в ПГА метод. При его использовании каждому элементу в популяции соответствует зона на колесе рулетки, пропорционально соразмерная с величиной ЦФ. Тогда при повороте колеса рулетки каждый элемент имеет некоторую вероятность выбора для селекции. Причём элемент с большим значением ЦФ имеет большую вероятность для выбора;
- селекция на основе заданной шкалы. Здесь популяция предварительно сортируется от «лучшей» к «худшей» на основе заданного критерия. Каждому элементу назначается определённое число, и тогда селекция выполняется согласно этому числу;

- элитная селекция. В этом случае выбираются лучшие (элитные) элементы на основе сравнения ЦФ. Далее они вступают в различные преобразования, после которых снова выбираются элитные элементы. Процесс продолжается аналогично до тех пор, пока продолжают появляться элитные элементы;
- турнирная селекция. При этом некоторое число элементов (согласно размеру «турнира») выбирается случайно или направленно из популяции, и лучшие элементы в этой группе на основе заданного турнира определяются для дальнейшего генетического поиска [4, 11–13];
- селекция на основе нечётких правил. Данный метод основан на применении элементов нечёткой логики. Ниже приведено правило для отбора решений:

IF ($Var = \text{мало}$) and ($G = \text{мало}$) then $p = \text{мало}$,

где p – вероятность выбора данной пары для операции кроссинговера;

$$Var = \frac{f_{\max} - \bar{f}}{f_{\max} - f_{\min}},$$

где $\bar{f}$ – среднее значение всех хромосом популяции; $f_{\max}$, $f_{\min}$ – лучшее и худшее значения функции пригодности; Var отражает степень различия функций пригодности.

G представляет расстояние между значением функции пригодности лучшей из хромосом-родителей и $f_{\max}$:

$$G = \frac{f_{\max} - \max(f(X), f(Y))}{f_{\max} - f_{\min}}.$$

Правило определяет, что две наиболее близкие хромосомы с лучшей функцией пригодности имеют большую вероятность подвергнуться кроссинговеру [14–17].

Оператор скрещивания (кроссовер).

Пусть $C_1 = (c_1^1, c_2^1, \dots, c_n^1)$ и $C_2 = (c_1^2, c_2^2, \dots, c_n^2)$ – две хромосомы, выбранные оператором селекции для скрещивания. Предполагается, что $c_k^1 < c_k^2$ и $f(C_1) > f(C_2)$.

Плоский кроссовер (flatcrossover): создаётся потомок $H = (h_1, \dots, h_k, \dots, h_n)$, h_k , $k = 1, \dots, n$ – случайное число из интервала $[c_k^1, c_k^2]$.

Простейший кроссовер (simplecrossover): случайным образом выбирается число k из интервала $\{1, 2, \dots, n-1\}$ и генерируются два потомка: $H_1 = (c_1^1, c_2^1, \dots, c_k^1, c_{k+1}^2, \dots, c_n^2)$ и $H_2 = (c_1^2, c_2^2, \dots, c_k^2, c_{k+1}^1, \dots, c_n^1)$. Схема простейшего кроссовера представлена на рисунке 6.

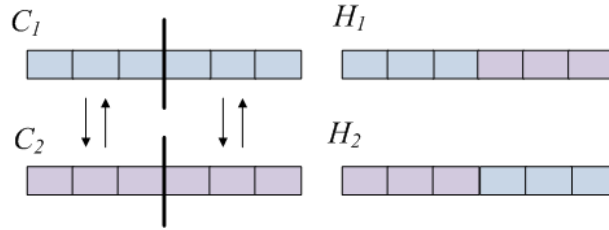


Рис. 6. Схема простейшего кроссовера

Двухточечный кроссовер. Аналогичен предыдущему, но случайным образом выбираются два числа k и t из интервала $\{1, 2, \dots, n-1\}$. Схема двухточечного кроссовера представлена на рисунке 7.

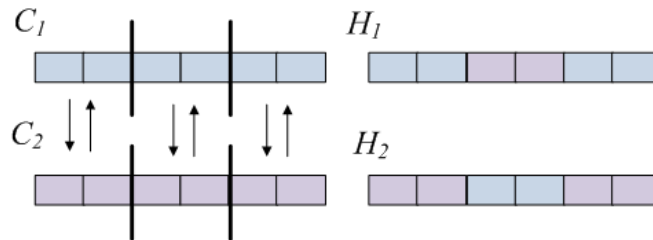


Рис. 7. Схема двухточечного кроссовера

Арифметический кроссовер (arithmeticalcrossover): создаются два потомка:

$$H_1 = (h_1^1, \dots, h_n^1), \quad H_2 = (h_1^2, \dots, h_n^2),$$

где $h_k^1 = w \cdot c_k^1 + (1-w) \cdot c_k^2$, $h_k^2 = w \cdot c_k^2 + (1-w) \cdot c_k^1$, $k=1, \dots, n$; w – либо константа (равномерный арифметический кроссовер) из интервала $[0; 1]$, либо изменяется с увеличением эпох (неравномерный арифметический кроссовер).

Геометрический кроссовер (geometricalcrossover): создаются два потомка:

$$H_1 = (h_1^1, \dots, h_n^1), \quad H_2 = (h_1^2, \dots, h_n^2),$$

где $h_k^1 = (c_k^1)^w \cdot (c_k^2)^{1-w}$, $h_k^2 = (c_k^2)^w \cdot (c_k^1)^{1-w}$, w – случайное число из интервала $[0; 1]$.

Смешанный кроссовер (blend, BLX-alpha crossover): генерируется один потомок:

$$H = (h_1, \dots, h_k, \dots, h_n),$$

где h_k – случайное число из интервала $[c_{\min} - I \cdot \alpha, c_{\max} + I \cdot \alpha]$, $c_{\min} = \min(c_k^1, c_k^2)$, $c_{\max} = \max(c_k^1, c_k^2)$, $I = c_{\max} - c_{\min}$. BLX=0,0 кроссовер превращается в плоский.

Линейный кроссовер (linearcrossover): создаются три потомка:

$$H_q = (h_1^q, \dots, h_k^q, \dots, h_n^q), \quad q=1, 2, 3,$$

где $h_k^1 = 0,5 \cdot c_k^1 + 0,5 \cdot c_k^2$, $h_k^2 = 1,5 \cdot c_k^1 - 0,5 \cdot c_k^2$, $h_k^3 = -0,5 \cdot c_k^1 + 1,5 \cdot c_k^2$.

На этапе селекции в этом кроссовере отбираются два наиболее сильных потомка.

Дискретный кроссовер (discretecrossover): каждый ген h_k выбирается случайно по равномерному закону из конечного множества $\{c_k^1, c_k^2\}$.

Расширенный линейчатый кроссовер (extendedlinecrossover): ген $h_k = c_k^1 + w \cdot (c_k^2 - c_k^1)$, w – случайное число из интервала $[-0,25; 1,25]$.

Эвристический кроссовер (Wright's heuristic crossover). Пусть C_1 – один из двух родителей с лучшей приспособленностью. Тогда $h_k = w \cdot (c_k^1 - c_k^2) + c_k^1$, w – случайное число из интервала $[0; 1]$.

Нечёткий кроссовер (fuzzyrecombination, FR-d crossover): создаются два потомка: $H_1 = (h_1^1, \dots, h_n^1)$, $H_2 = (h_1^2, \dots, h_n^2)$. Вероятность того, что в i -м гене появится число v_i , задаётся распределением $p(v_i) \{F(c_k^1), F(c_k^2)\}$, где $F(c_k^1)$, $F(c_k^2)$ – распределения вероятностей треугольной формы (треугольные нечёткие функции принадлежности) со следующими свойствами: $(c_k^1 \leq c_k^2 \text{ и } I = |c_k^1 - c_k^2|)$ [11, 18–24].

Комбинированный кроссовер. Данный тип оператора был разработан автором. В результате работы кроссовера генерируется один потомок. Случайным образом выбирается число точек разрыва k из интервала $\{1, 2, \dots, n-1\}$, а также задаётся количество родителей n . Если $n = k+1$, то потомок генерируется следующим образом: из каждой хромосомы-родителя выбирается блок согласно её номеру в массиве родителей. Схема комбинированного кроссовера для данного случая представлена на рисунке 8.

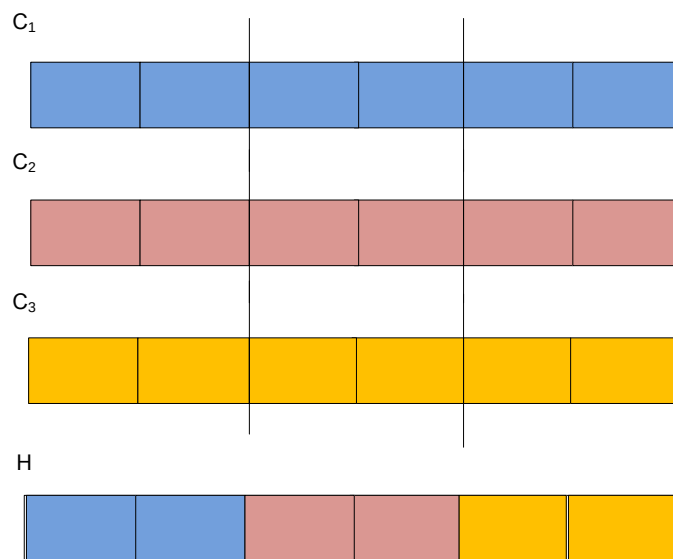
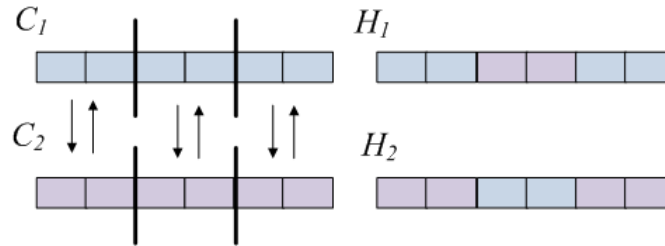


Рис. 8. Схема комбинированного кроссовера для случая $n = k+1$

Если $n < k+1$, то потомок генерируется следующим образом: из каждой хромосомы-родителя выбирается блок согласно её номеру в массиве родителей, после чего процесс начинается сначала, но блок выбирается согласно номеру хромосомы плюс 1 и т.д., пока не будет сформирован потомок. Схема комбинированного кроссовера для данного случая представлена на рисунке 9.

Рис. 9. Схема комбинированного кроссовера для случая $n < k+1$

Оператор мутации совместно с кроссинговером обеспечивает разнообразие в популяции. Классическая битовая мутация не подходит для НГА. Рассмотрим возможные реализации оператора мутации, для этого введём следующие обозначения:

- $C = (c_1, c_2, \dots, c_n)$ есть мутируемая хромосома;
- $c_i \in [a, b]$ – мутируемый ген на поколении t ;
- $c_i' \in [a, b]$ – этот же ген после мутации.

Одноточечная мутация. Случайно выбирается ген в хромосоме и случайным образом обменивает его на рядом расположенный ген. Схема одноточечной мутации представлена на рисунке 10.

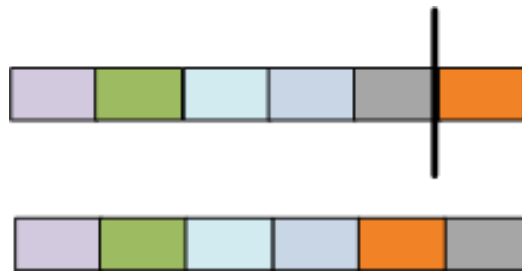


Рис. 10. Схема одноточечной мутации

Двухточечный ОМ заключается в перестановке генов, расположенных справа от точек разрыва.

Простой оператор мутации. Случайно выбирается ген в хромосоме и случайным образом изменяется. Схема простого оператора мутации представлена на рисунке 11.

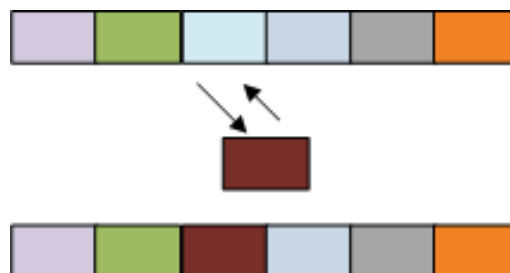


Рис. 11. Схема простого оператора мутации

Вещественная мутация:

$c_i' = c_i + \chi \cdot r \cdot (b_i - c_i)$, если $\chi = 0$, или $c_i' = c_i - \chi \cdot r \cdot (c_i - a_i)$, если $\chi = 1$, где $\chi \in \{0, 1\}$ – случайная величина, выбираемая при каждой мутации, которая определяет направление мутации; $r \in [0, 1]$ – случайная величина, выбираемая при каждой мутации, которая определяет силу мутации.

Вещественная мутация с элементами имитации отжига:

$c_i' = c_i + \chi \cdot r \cdot (b_i - c_i)$, если $\chi = +1$, или $c_i' = c_i + \chi \cdot r \cdot (c_i - a_i)$, если $\chi = -1$, где $\chi \in \{-1, 1\}$ – случайная величина, выбираемая при каждой мутации, которая определяет направление мутации; $r = r_0 \cdot \exp(1/s)$, $r, r_0 \in [0, 1]$ – величина, определяющая силу мутации (r_0 – величина мутации, определяемая перед запуском алгоритма и указывающая максимальную силу мутации); s – множитель, определяющий степень изменения мутации в зависимости от поколения [4, 8, 11].

Инверсия и сегрегация. В операторе инверсии случайным образом определяется одна или несколько точек инверсии, внутри которых элементы инвертируются. На рисунке 12 представлена схема оператора инверсии.



Рис. 12. Схема оператора инверсии

Оператор сегрегации реализуется на некотором наборе хромосом. Пусть имеется популяция P , состоящая из четырёх родительских хромосом $P = \{P_1, P_2, P_3, P_4\}$: $P_1 : (1234)$; $P_2 : (2431)$; $P_3 : (3142)$; $P_4 : (4321)$. Тогда потомок P'_1 можно сформировать случайным образом, взяв первый элемент (один или несколько генов) из P_1 , второй – из P_2 , третий – из P_3 и четвёртый – из P_4 . При этом должны быть отброшены повторяющиеся решения или решения, содержащие одинаковые элементы. Так, вариант $P'_1 : (1412)$ должен быть отброшен как содержащий две единицы, а вариант $P'_1 : (2143)$ является легальным (допустимым или реальным) [4, 6].

Как было сказано ранее, для повышения эффективности проектируемого генетического алгоритма необходимо правильно сочетать различные виды генетических операторов, причём необходимо также применение механизмов, способных в нужные моменты, анализируя состояния популяции, менять параметры генетических операторов или изменять сам набор операторов в алгоритме. Прорывом на новый уровень может служить разработка новых генетических операторов с применением нечёткой логики или других парадигм.

3. Проблемы при применении генетического алгоритма и методы их решений

При использовании генетического алгоритма предстоит столкнуться с большим количеством трудностей, от решения которых и будет зависеть эффективность алгоритма. Ниже будут приведены наиболее часто встречающиеся проблемы и возможные методы их решения.

Подбор параметров генетического алгоритма. Набор параметров генетического алгоритма можно представить в виде вектора $\vec{X} = \{x_1, x_2, \dots, x_n\}$. Именно от правильного подбора параметров зависит эффективность алгоритма. Поэтому необходима разработка методов подбора параметров и изменения их в процессе развития популяции. Для этого необходимо определить, какой из параметров вносит наибольший вклад путём небольших изменений каждого из параметров и последующей оценки. После чего в зависимости от развития популяции происходит изменение одного или нескольких параметров.

Генетический алгоритм имеет преимущество перед стандартными методами только на сложных и многокритериальных задачах. Поэтому одной из проблем является определение возможности решения поставленной задачи с помощью ГА. Следовательно, необходимо тщательно изучить поставленную задачу, например, изучить ландшафт целевой функции путём генерации чисел и определения её значений. По этим значениям можно определить примерный ландшафт функции, и, если он достаточно прост, задачу лучше решать стандартными общепринятыми методами.

Большое значение для генетического алгоритма имеет генерация начальных решений. Необходимо правильно выбрать стратегию генерации начальной популяции, чтобы она как можно более полно охватывала область решений. Для этого необходимо тщательно изучить область допустимых решений. Но также важно отслеживать и в процессе работы алгоритма подобласть, занимаемую хромосомами как одной популяции, так и всех. Решением данной проблемы может служить разбиение исходной области на более мелкие подобласти и исследование каждой области отдельной популяцией или несколькими популяциями.

Поиск нескольких оптимумов. Для решения данной проблемы необходимо применение многопопуляционных алгоритмов, причём каждая популяция должна исследовать свой участок пространства решения.

Преждевременная сходимость алгоритма. Для решения этой проблемы необходимо использование методов, позволяющих изменять параметры генетического алгоритма в процессе развития популяции. Эффективным методом решения данной проблемы также является применение многопопуляционного алгоритма с использованием миграции особей между популяциями.

Заключение

Генетические алгоритмы, благодаря своей простоте, получили широкое распространение. Но эффективность генетического алгоритма сильно зависит от выбранных параметров. Зачастую небольшое изменение одного из них может привести к резкому изменению результата. Как было показано в статье, для эффективного применения генетических алгоритмов необходимо провести предварительно большую аналитическую работу. Следует помнить, что применение ГА полезно лишь в тех случаях, когда для данной задачи нет подходящего специального алгоритма решения. За время своего развития генетические алгоритмы прошли большой путь развития от довольно простых алгоритмов с небольшим набором операторов до довольно сложных, сочетающих в себе разные парадигмы. Но путь развития не останавливается, а продолжается. Необходима разработка алгоритмов, одинаково хорошо способных решать различные задачи, при этом способных менять свою структуру в зависимости от поставленных задач. Также необходима более тесная интеграция различных эволюционных парадигм, ведущих к созданию более эффективных алгоритмов и способных вывести решение задач на качественно новый уровень.

Список литературы

1. Ярушкина Н.Г. Основы теории нечётких и гибридных систем. – М.: Финансы и статистика, 2004 – С. 324.
2. Гладков Л.А., Курейчик В.В., Курейчик В.М. Генетические алгоритмы / под ред. В.М. Курейчика. – М.: Физматлит, 2006.
3. Хедрик Ф. Генетика популяций. – М.: Техносфера, 2003.
4. Курейчик В.М., Кныш Д.С. Параллельный генетический алгоритм модели и проблемы построения // Интегрированные модели и мягкие вычисления в искусственном интеллекте: сб. науч. тр. V Междунар. науч.-практич. конф. – М.: Физматлит, 2009. – С. 41–51.
5. Емельянов В.В., Курейчик В.В., Курейчик В.М. Теория и практика эволюционного моделирования. – М.: Физматлит, 2003.
6. Джонс М.Т. Программирование искусственного интеллекта / пер. с англ. А.И. Осипова. – М.: ДМК Пресс, 2004. – 312 с.: ил.
7. Панченко В.Т. Генетические алгоритмы. – Астрахань: Астраханский университет, 2007.
8. Голубин А.В. Разработка гибридных интеллектуальных моделей эволюционного проектирования: дис. ... канд. техн. наук. – М., 2006.
9. Abdullah S., Turabieh H. Generating University Course Timetable Using Genetic Algorithms and Local Search // In the Third 2008 International Conference on Convergence and Hybrid Information Technology ICCIT. – 2008. – Vol. I. – P. 254–260.
10. Whitley D., Garrett D., Watson J.P. Quad Search and Hybrid Genetic Algorithms // Genetic and Evolutionary Computation Conference (GECCO 2003). – Pp. 1469–1480.

11. Непрерывные генетические алгоритмы – математический аппарат. – URL: http://www.basegroup.ru/library/optimization/real_coded_ga.
12. Oates M.J., Corne D.W. & Kell, D.B. (2003). The bimodal feature at large population sizes and high selection pressure: implications for directed evolution. In Recent advances in simulated evolution and learning (ed. K.C. Tan, M.H. Lim, X. Yao and L. Wang). – World Scientific, Singapore.
13. Herrera F., Lozano M., Verdegay J.L. Tackling real-coded Genetic algorithms: operators and tools for the behaviour analysis // Artificial Intelligence Review. – 1998. – Vol. 12, No. 4. – P. 265–319.
14. Нечётные гибридные системы. Теория и практика / И.З. Батыршин [и др.]; под ред. Н.Г. Ярушкиной. – М.: Физматлит, 2007.
15. Батыршин И.З. Основные операции нечёткой логики и их обобщения. – Казань: Отечество, 2001. – 100 с.: ил.
16. Кофман А. Введение в теорию нечётких множеств. – М.: Радио и связь, 1982. – 432 с.
17. Тэрано Т., Асаи К., Сугено М. Прикладные нечёткие системы. – М.: Мир, 1993 – 368 с.
18. Herrera F., Lozano M., Sanchez A.M. Hybrid Crossover Operators for Real-Coded Genetic Algorithms: An Experimental Study // Soft Comput. 9(4): 280-298 (2005).
19. Wright A. Genetic algorithms for real parameter optimization // Foundations of Genetic Algorithms. – 1991. – Vol. 1. – P. 205–218.
20. Corne D.W., Oates M.J. & Kell D.B. (2002) On fitness distributions and expected fitness gain of mutation rates in parallel evolutionary algorithms. In Parallel Problem Solving from Nature – PPSN VII. (eds. J.J. Merelo Guervós, P. Adamidis, H.-G. Beyer, J.-L. Fernández-Villacañás & H.-P. Schwefel) 132-141 (Springer, Berlin; 2002).
21. Herrera F., Lozano M., Verdegay J.L. Tackling real-coded Genetic algorithms: operators and tools for the behaviour analysis // Artificial Intelligence Review. – 1998. – Vol. 12, No. 4. – P. 265–319.
22. Herrera F., Lozano M., Sanchez A.M. Hybrid Crossover Operators for Real-Coded Genetic Algorithms: An Experimental Study // Soft Comput. 9(4): 280-298 (2005).
23. Deb K. and Kumar A. (1995). Realcoded genetic algorithms with simulated binary crossover: Studies on multimodal and multiobjective problems. ComplexSystems, 9(6), 431–454.
24. Wright A. Genetic algorithms for real parameter optimization // Foundations of Genetic Algorithms. – 1991. – Vol. 1. – P. 205–218.
25. Corne D.W., Oates M.J. & Kell D.B. (2003). Fitness gains and mutation patterns: deriving mutation rates by exploiting landscape data. In Foundations of Genetic Algorithms 7 (ed. K. A. De Jong, R. Poli and J. E. Rowe), – Pp. 347–364. Morgan Kaufmann, Amsterdam.

УДК 004.386: 004.896

М.В. Ляшов

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ КОДИРОВАНИЯ СОСТОЯНИЙ КОНЕЧНОГО АВТОМАТА

Введение. Любая цифровая система состоит из комбинационных и последовательностных логических схем, при этом поведение последовательностной схемы может быть описано конечным автоматом (КА). Методы синтеза КА непрерывно развиваются [1–5]. Один из самых важных шагов в синтезе КА – это кодирование внутренних состояний. Задача кодирования внутренних состояний состоит в обнаружении лучшего назначения комбинаций триггеров к состояниям автомата. При минимальном числе переключений элементов памяти уменьшается сложность схем, реализующих дизъюнкции на входах элементов памяти, что приводит к минимизации комбинационной схемы автомата [1].

В современных системах автоматизированного проектирования для кодирования внутренних состояний конечного автомата используются алгоритмы NOVA, MUSTANG и их модификации. Однако сгенерированные коды состояний для КА являются далёкими от оптимума [2]. В данной статье рассмотрен разработанный эволюционный алгоритм кодирования внутренних состояний КА, показавший лучшие результаты по сравнению с традиционными методами.

Задача синтеза конечных автоматов. Традиционно процесс проектирования конечного автомата состоит из четырёх шагов (рис. 1):

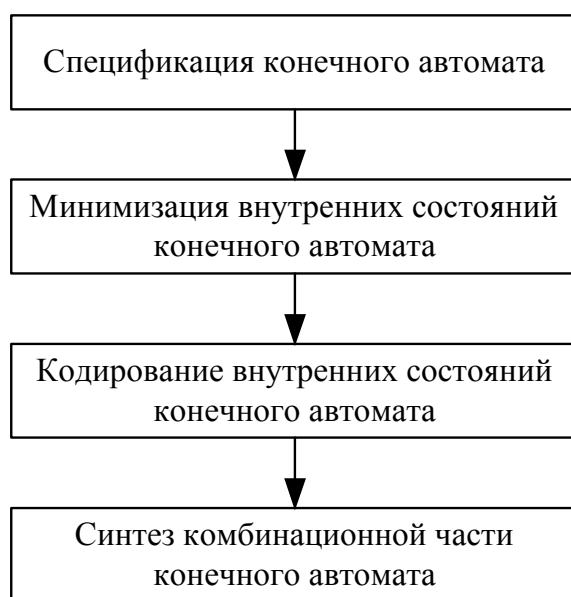


Рис. 1. Процесс проектирования конечного автомата

1. Спецификация последовательстной системы, которая должна определить следующие состояния и выходы каждого существующего состояния автомата. Это делается путём использования диаграммы или таблицы состояний.

2. Минимизация внутренних состояний призвана уменьшить число состояний конечного автомата, используя эквивалентность и т.п.

3. Кодирование внутренних состояний должно назначить отличную комбинацию на каждое состояние автомата.

4. На этапе синтеза комбинационной части конечного автомата, когда используются классические методы синтеза и минимизации, реализуется схема управления.

Как только спецификация и шаг минимизации состояний были закончены, выполняется следующий шаг, состоящий из назначения кода к каждому состоянию. Если автомат имеет N различных состояний, тогда требуется N различных комбинаций двоичного кода. Следовательно, требуется K триггеров, чтобы хранить текущее состояние, где K – наименьшее положительное целое число, такое, что $2^K \geq N$. Проблема кодирования внутренних состояний состоит в обнаружении лучшего назначения комбинаций триггеров к состояниям автомата. Так как состояние автомата – только устройство подсчёта, комбинационная логика управления необходима для активации триггеров в желательной последовательности.

Компонент комбинационной логики управления в конечном автомате ответствен как за генерацию первичных сигналов на выходах, так и за генерацию сигналов для переключения состояний. Традиционно комбинационная логика управления создаётся, используя таблицы переходов триггеров. Сложность логики управления различна для различных комбинаций триггеров. При минимальном числе переключений элементов памяти уменьшается сложность схем, реализующих дизъюнкции на входах элементов памяти, что приводит к минимизации комбинационной схемы автомата [1]. Следовательно, проектировщик должен искать назначение, которое минимизирует сложность комбинационной логики.

Все методы синтеза конечных автоматов делятся на классические, применимые к любому конечному автомату независимо от способа его реализации; двухуровневый синтез, ориентированный для реализации на PLD со структурой двух программируемых матриц И и ИЛИ; и многоуровневый синтез, используемый, в основном, при реализации конечных автоматов на FPGA. А также методы, не вошедшие в предыдущие классы, и методы синтеза, явно ориентированные для реализации конечных автоматов на PLD и позволяющие учитывать архитектурные особенности последних [2].

Число различных и неравнозначных вариантов кодирования для конкретного автомата определяется выражением [1]:

$$C = \frac{(2^K - 1)!}{(2^K - N)!K!}, \quad (1)$$

где N – число состояний автомата; K – минимальная длина булева вектора, кодирующего состояние.

Очевидно, что перебрать все возможные варианты реально только при небольшом числе состояний автомата. Если же число состояний достаточно велико, то полный перебор неэффективен и для выбора не самого плохого варианта кодирования необходимо применять другие методы.

Генетический алгоритм кодирования состояний. Целевая функция

При решении задачи кодирования состояний конечного автомата в алгоритмах минимизации числа переключений используется целевая функция:

$$F = \sum_{k=0}^p |Xq_i, Xq_j|_k, \quad (2)$$

где $|Xq_i, Xq_j|_k$ – расстояние по Хеммингу между кодами состояний k -го перехода, p – количество переходов. Весовая функция определяет сложность (аппаратные затраты) на построение комбинационной схемы функции возбуждения, т.е. является одной из оценок сложности комбинационной схемы автомата [1–3].

Для вычисления целевой функции необходимо граф либо таблицу переходов КА преобразовать в матрицу смежности. Строки соответствуют текущему состоянию автомата, а столбцы – состоянию, в которое автомат переходит при поступлении на его вход соответствующего сигнала. Пример преобразования таблицы переходов представлен на рисунке 2.

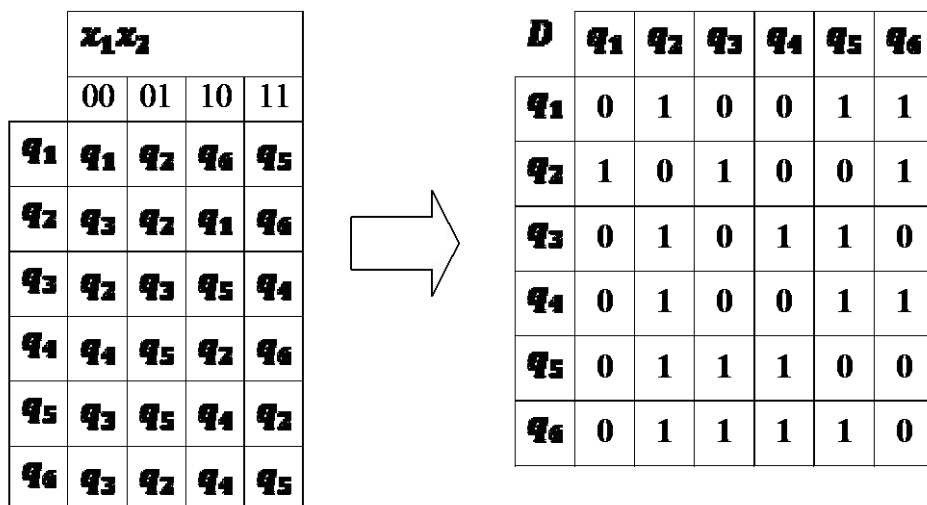


Рис. 2. Пример преобразования таблицы переходов в матрицу смежности

Используя матрицу смежности, целевую функцию (2) можно представить в следующем виде:

$$F = \sum_{i=1}^N \sum_{j=1}^N |Xq_i, Xq_j| \cdot D_{ij}, \quad (3)$$

где N – число состояний конечного автомата.

Количество переходов при этом будет равно:

$$p = \sum_{i=1}^N \sum_{j=1}^N D_{ij}. \quad (4)$$

Задачей генетического алгоритма является минимизация целевой функции:

$$F \rightarrow \min.$$

Предельным случаем задачи является кодирование, при котором $F \rightarrow p$, т.е. одно переключение элемента памяти на каждом переходе.

Генетические операторы и структура генетического алгоритма

Основными генетическими операторами, используемыми при работе генетического алгоритма, являются операторы кроссинговера и мутации. Использование различных методов представления хромосом накладывает ограничения на типы применяемых операторов. Оператор кроссинговера должен сохранять как можно больше информации от исходных хромосом, создавая новые. Согласно целевой функции расстояние по Хэммингу между двумя битовыми строками – это расстояние между отдельными битами в строке, т.е. целевую функцию можно представить как сумму колонок двоичной матрицы. Таким образом, в результате применения оператора кроссинговера в новой хромосоме должно присутствовать больше колонок от хромосомы с лучшей целевой функцией. Поэтому для определения точки разрыва хромосомы вводится весовой коэффициент, зависящий от значения целевой функции.

Рассмотрим работу этого кроссинговера, изображённого на рисунке 3, на примере конечного автомата в шести состояниях.

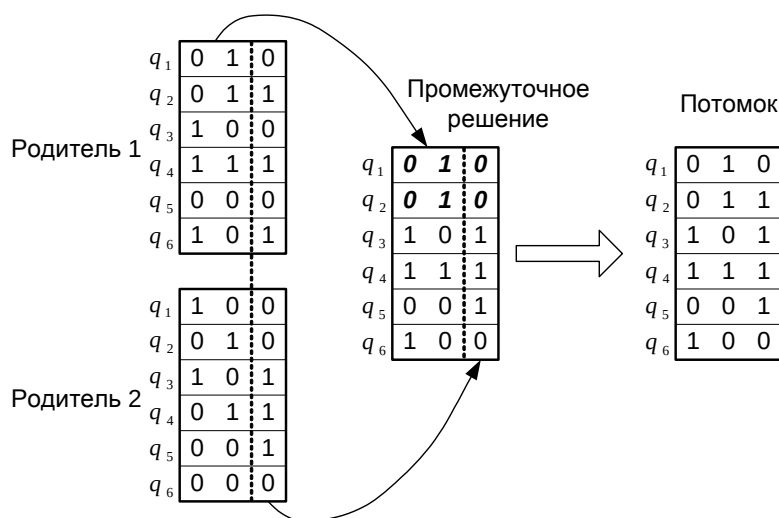


Рис. 3. Пример работы оператора кроссинговера

Пусть Родитель 1 имеет лучшую целевую функцию по сравнению с Родитель 2, точкой разрыва хромосом будет колонка 2. Получаем промежуточное решение, которое должно быть проверено на наличие нелегальных кодов состояний. Нелегальной считается такая кодировка, у которой разные состояния имеют одинаковые двоичные коды. В приведённом примере состояния q_1 и q_2 имеют одинаковый двоичный код 010. Чтобы исправить нелегальное решение, один из кодов заменяется свободным двоичным кодом из множества 2^K .

В результате кроссинговера получаем одного потомка. Основное преимущество данного оператора кроссинговера для решения задачи назначения состояний КА по сравнению с операторами, представленными в работах [3–5] в том, что после его применения в популяции сохраняется информация от хромосом с лучшей целевой функцией.

Так как оператор кроссинговера сохраняет информацию от хромосом с лучшей целевой функцией в популяции, то поиск будет охватывать ограниченное пространство всевозможных решений, что, в свою очередь, может снизить вероятность нахождения лучшего решения. Для создания в популяции случайной информации необходим оператор мутации.

Пример выполнения оператора мутации показан на рисунке 4. Для выполнения операции мутации случайным образом в хромосоме выбираются два кода состояний и меняются местами.

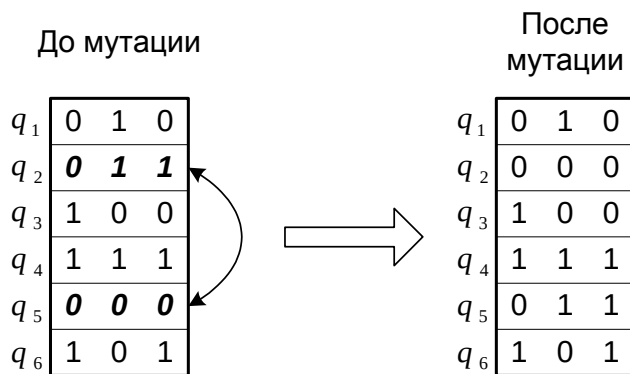


Рис. 4. Пример работы оператора мутации

Выбор оператора отбора зависит от типа применяемого генетического алгоритма. Для решения задачи назначения состояний КА выбран генетический алгоритм устойчивого состояния. Поскольку после оператора кроссинговера и оператора мутации к популяции нужно добавить 2 хромосомы, то оператор отбора должен удалить 2 хромосомы. Поэтому необходимо устранять дублирующие решения из популяции. Оператор отбора будет состоять из двух этапов. Сначала ищутся и удаляются две дублирующие хромосомы, а затем, если было удалено менее двух хромосом, делается просмотр в порядке возрастания фитнеса и удаление с вероятностью P текущего решения.

Структурная схема генетического алгоритма назначения состояний конечного автомата представлена на рисунке 5. Приведённая схема генетического алгоритма обладает операциями, учитывающими специфику решаемой задачи. Так, методики кодирования и декодирования хромосомы используют графы недопустимых кодировок, что препятствует возникновению «нелегальных» решений. Эти особенности позволяют уменьшить пространство поиска и тем самым ускорить получение качественных решений.

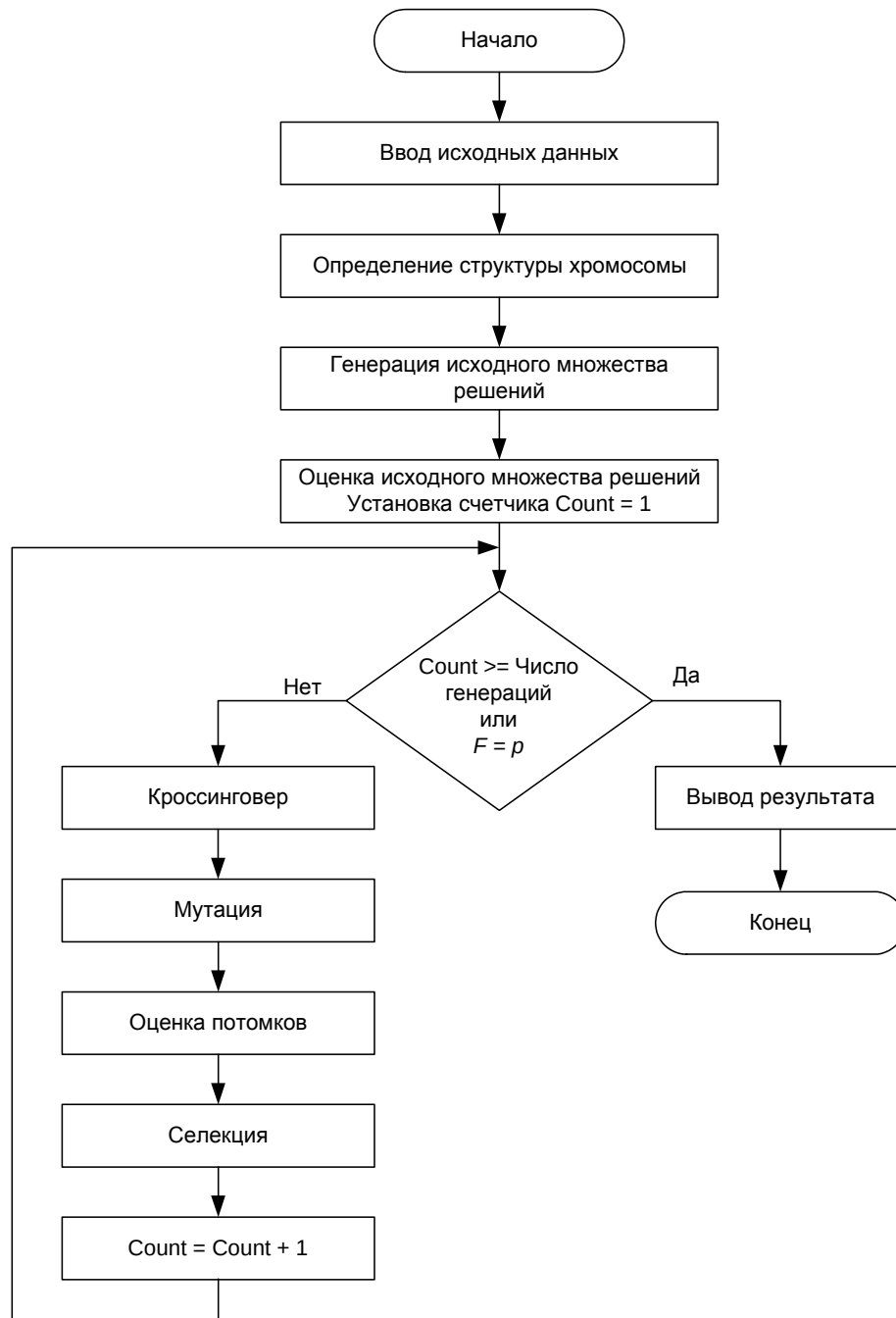


Рис. 5. Структурная схема генетического алгоритма назначения состояний конечного автомата

Сравнение результатов. На рисунке 6 приведено сравнение эффективности разработанного генетического алгоритма с двумя вариантами алгоритма NOVA и аналогичного ГА, приведённого в работе [3].

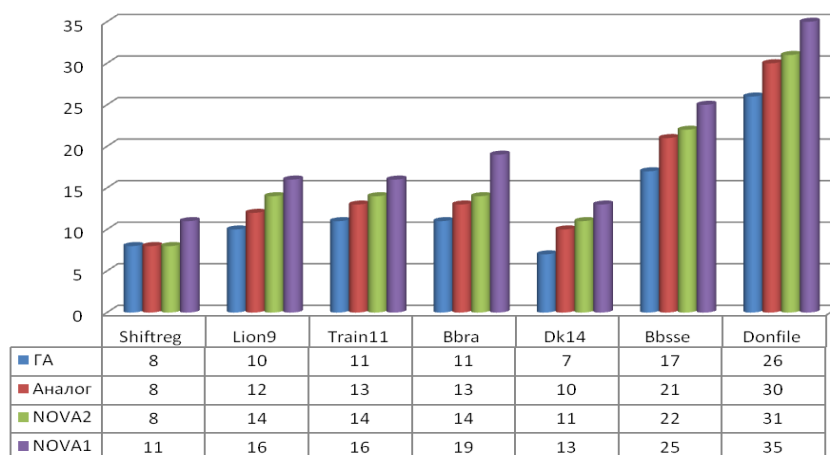


Рис. 6. Сравнение эффективности различных алгоритмов для тестовых КА

Из графика видно, что разработанный генетический алгоритм на всех тестовых КА имеет в среднем на 15–33 % меньшее значение целевой функции. Таким образом, разработанный генетический алгоритм назначения состояний синхронного конечного автомата показывает результаты, которые превосходят классические методы и аналоги, что позволяет его использовать в современных САПР.

Список литературы

1. Закревский А.Д., Поттосин Ю.В., Черемисинова Л.Д. Логические основы проектирования дискретных устройств. – М.: ФИЗМАТЛИТ, 2007. – 592 с.
2. Соловьев В.В. Логическое проектирование цифровых систем на основе программируемых логических интегральных схем. – М.: Горячая Линия – Телеком, 2008. – 376 с.
3. Nedjah N. Evolvable Machines: Theory and Practice. Studies in Fuzziness and Soft Computing. – 2004.– Vol. 161. – 260 с.
4. Chyz'y M., Kosin'ski W. Genetic Algorithm for the State Assignment Problem // Communications of the 10th International Symposium Intelligent Information Systems, Zakopane, Poland, 17–21 June 2001. – Pp. 7–11.
5. Belgasem A., Kalganova T., Almaini A. (2002). Extrinsic Evolution of Finite State Machine. ACDM2002, UK. I. C. Paimee (Ed.) Published by Springer. – Pp. 157–168.
6. Гладков Л.А., Курейчик В.В. Генетические алгоритмы. – М.: Физматлит, 2006. – 320 с.

УДК 517.11

А.В. Козоброд, В.Е. Мешков, Е.В. Мешкова

ОБРАБОТКА ТЕКСТОВЫХ ДОКУМЕНТОВ НА ОСНОВЕ ЗАКОНОВ ЗИПФА ДЛЯ ФОРМИРОВАНИЯ ГИБРИДНОГО НЕЙРОСЕТЕВОГО КЛАССИФИКАТОРА

В данной статье изложены принципы применения законов Зипфа для предварительной обработки текстовых документов, используемых для формирования гибридного нейросетевого классификатора. Рассмотрена программа, реализующая данный подход, и её возможности для обработки и анализа текстовых документов.

Ключевые слова: Зипф, текстовая информация, нейросетевой, классификация, анализ.

Большинство современных поисковых систем используют методы поиска информации, основанные на двух простых концепциях. На входе системы «запрос» представляет собой последовательность знаков, описывающих искомую информацию. В этом случае каждое слово запроса называется «поисковым термином». Запрос может состоять из одного поискового термина, фразы на естественном языке или из сложного выражения с использованием специальных символов.

В первом случае методы поиска сравнивают запрос с каждым словом в тексте, не делая разницы между функциями различных слов (это так называемый «полнотекстовый поиск» [1]).

Во втором случае поисковые методы идентифицируют поля библиографической или структурной информации (например, автор или заголовок) и проводят поиск только в этих полях (это называют «поиском по полям» [1]).

Оба эти метода достаточно эффективны, поэтому современные системы обработки информации часто используют их в некоторой комбинации [1]. Если поиск по полям требует чётких методов определения полей, то полнотекстовый поиск может быть эффективен даже в необработанном тексте. Однако разнородные тексты переменной длины, стиля и содержания классифицировать таким методом затруднительно, а результаты могут не воспроизводиться или быть неадекватными. Вследствие этой причины используются методы поиска с логическими операторами и ранжирование ответов. Но это не решает всех проблем. Алгоритмы ранжирования дают мало информации о том, на каком основании принимается решение о расположении записей в списке. В результате наивысший ранг может быть присвоен странице с минимальной ценностью для пользователя, а действительно полезные страницы будут далеко внизу списка [1].

В связи с этим возникает необходимость в новых методах поиска информации, которые были бы лишены перечисленных выше недостатков.

Этими свойствами обладают системы, построенные на основе ассоциативных нейронных сетей [2], которые позволяют находить информацию, полностью удовлетворяющую запросу.

При использовании нейронных сетей при построении систем для обработки классификации текстовой информации одной из трудностей является ориентация нейронных сетей на обработку сигналов, а не символьной информации. В наиболее развитых моделях искусственных нейронных сетей входная информация представляется в виде двоичного вектора. При этом возникает проблема кодирования входной информации и декодирования выходного вектора. Если кодировать символьную информацию на входе искусственной нейронной сети бессистемно, т.е. не заботиться о взаимосвязи между значениями двоичных векторов и соответствующими им символьными значениями, то близкие по смыслу символьные значения могут кодироваться совершенно разными двоичными векторами, отстоящими друг от друга на очень большое расстояние. Это усложняет обучение нейронной сети и может приводить к ошибкам при функционировании нейронной сети [3].

На выходе нейронной сети полученный в результате работы нейронной сети двоичный вектор необходимо декодировать, т.е. преобразовывать его в одно или несколько символьных значений. При декодировании выходного вектора надо учитывать одно из возможных требований, которые могут предъявляться к решению, получаемому нейронной сетью:

- исключить или уменьшить вероятность ложного решения;
- исключить отсутствие какого-либо решения, например, за счёт качества.

В первом случае необходимо вводить избыточность в кодирование значений символьных решений и двоичные векторы, не соответствующие кодам значений из тезауруса решений, не декодировать (случай «отсутствия решения»). Во втором случае избыточность при кодировании не требуется, а если она используется (например, для обеспечения симметричности методов кодирования на входе и выходе сети), в случае несоответствия выходного вектора какому-либо из значений тезауруса выбирается и декодируется ближайший по значению код.

Для исключения этих недостатков при использовании нейронных сетей для классификации символьной информации предлагается использовать предварительную обработку текстового потока на основе законов Зипфа [3, 4]. С учётом первого и второго законов Зипфа [1] все созданные человеком тексты построены по единым правилам, т.е. какой бы язык ни использовался, кто бы ни написал данный текст, внутренняя структура текста останется неизменной.

Предложенный подход представляется возможным использовать при создании системы поиска и классификации текстовой информации с использованием нейронных сетей.

Первый закон Зипфа – «ранг – частота»

Для любого слова из анализируемого текста определим частоту вхождения слова в текст. Отметим, что некоторые слова могут иметь одинаковую частоту, т.е. входить в текст равное количество раз. Сгруппируем их, взяв только одно значение из каждой группы. Расположим частоты по мере их убывания и пронумеруем (индексируем). Порядковый номер частоты называется рангом частоты. Вероятность появления слова в данном тексте будет равна отношению частоты вхождения этого слова к общему числу слов в тексте:

$$\text{Вероятность} = \text{Частота вхождения слова} / \text{Число слов.}$$

Согласно первому закону Зипфа, если умножить вероятность обнаружения слова в тексте на ранг частоты, то получившаяся величина (C) приблизительно постоянна:

$$C = (\text{Частота вхождения слова} \times \text{Ранг частоты}) / \text{Число слов.}$$

Очевидно, что это функция типа $y=k/x$ и её график – равносторонняя гиперболоа.

Значение константы в разных языках различно, но внутри одной языковой группы остаётся неизменной, какой бы текст ни исследовался. Так, например, для английских текстов константа Зипфа равна приблизительно 0,1. Для русского языка коэффициент Зипфа равен 0,06-0,07.

Второй закон Зипфа – «количество – частота»

По первому закону Зипфа получается, что разные слова могут входить в текст с одинаковой частотой. Зипф установил, что частота и количество слов, входящих в текст с этой частотой, тоже связаны между собой. Если построить график, отложив по одной оси (оси X) частоту вхождения слова, а по другой (оси Y) – количество слов в данной частоте, то получившаяся кривая будет сохранять свои параметры для всех без исключения созданных человеком текстов. Как и в предыдущем случае, это утверждение верно в пределах одного языка. Однако и межъязыковые различия невелики. На каком бы языке текст ни был написан, форма кривой Зипфа останется неизменной. Могут немного отличаться лишь коэффициенты, отвечающие за наклон кривой (в логарифмическом масштабе, за исключением нескольких начальных точек, график – прямая линия). Следовательно, можно утверждать, что законы Зипфа универсальны. В принципе, они применимы не только к текстам. В аналогичную форму выливается, например, зависимость количества городов от числа проживающих в них жителей. Характеристики популярности узлов в сети Интернет тоже отвечают законам Зипфа. Не исключено, что в законах отражается «человеческое» происхождение объекта [1].

Программа анализа текста на основе законов Зипфа

Рассмотрим практическую реализацию данного подхода, т.е. программу обработки произвольного текста и выделения из него «ключевых слов Зипфа». Интерфейс программы анализа текста на основе законов Зипфа показан на рисунке 1.

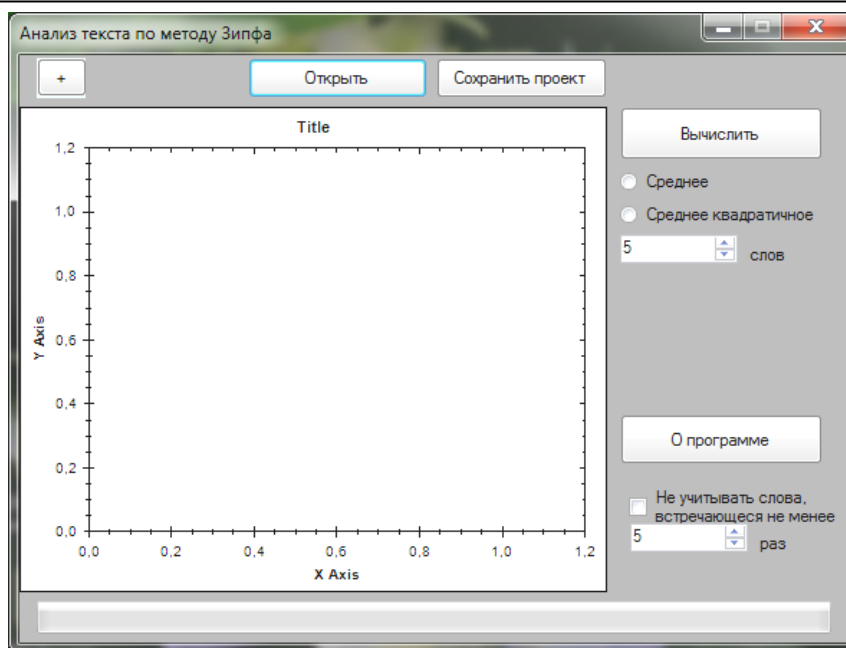


Рис. 1. Интерфейс программы анализа текста

Разработанная программа позволяет открыть произвольный текст в форматах *.txt, *.rtf, *.html или файл с произвольного типа, но в текстовом формате (режим «ОТКРЫТЬ»). Обработанный файл (проект) возможно сохранить для дальнейшего анализа (режим «СОХРАНИТЬ ПРОЕКТ»). Другие режимы работы проиллюстрируем в ходе описания работы над проектом (обработки файла).

Для файла формата *.html, посвящённого законодательной сфере в области защиты прав потребителей, воспользуемся первым законом Зипфа и построим график зависимости ранга от частоты. Как уже упоминалось, его форма всегда одинакова. Пример анализа упомянутого файла текста показан на рисунке 2.

Исследования показывают, что наиболее значимые слова лежат в средней части диаграммы. Слова, которые попадают слишком часто, в основном оказываются предлогами, местоимениями, в английском – артиклями и т.п. Редко встречающиеся слова тоже в большинстве случаев не имеют решающего смыслового значения. От того, как будет определён диапазон значимых слов, зависит отражение структуры текста на нейронной сети. На примере (рис. 2) для обработки документа выбраны следующие режимы:

- не учитываются слова, встречающиеся не менее 10 раз;
- вычисляется среднеквадратичное значение (возможно вычисление среднего).

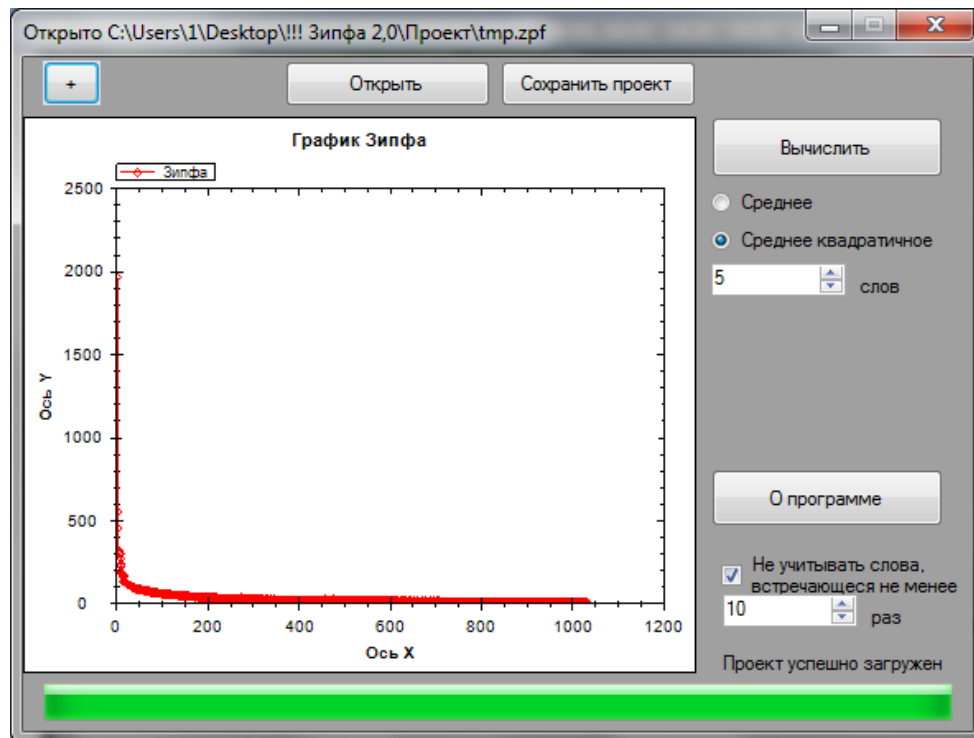


Рис. 2. Пример анализа текстового документа

В режиме «ПРОМЕЖУТОК» (рис. 3) имеется возможность управления интервалом учитываемых и отображаемых на графике слов.

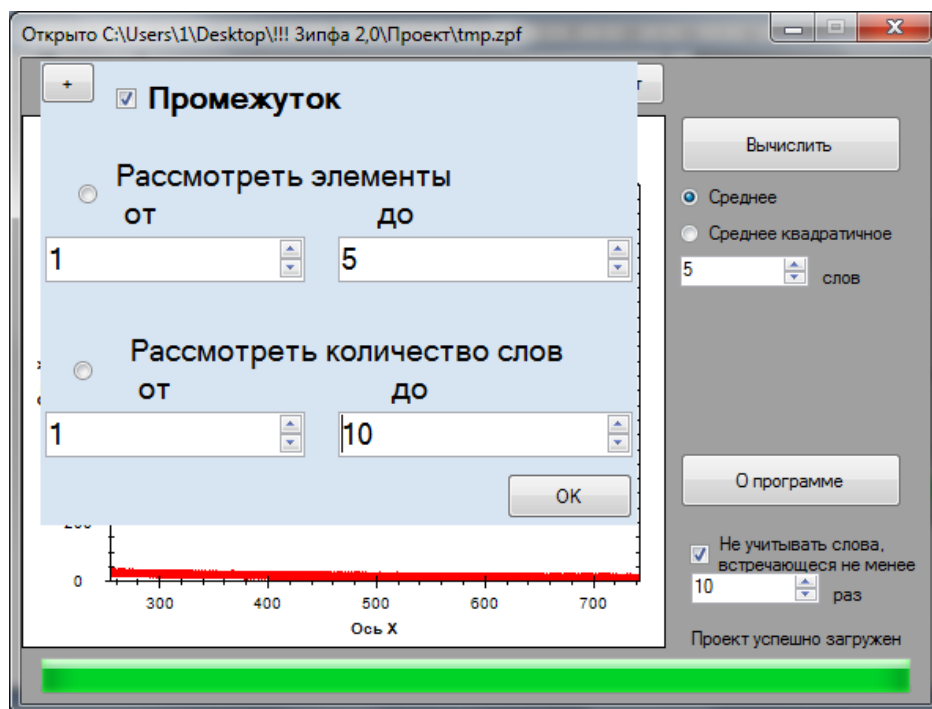


Рис. 3. Управление интервалами анализа

Параметры, установленные для интервалов анализа текста (рис. 3), приводят к следующей коррекции графика гиперболы (рис. 4).

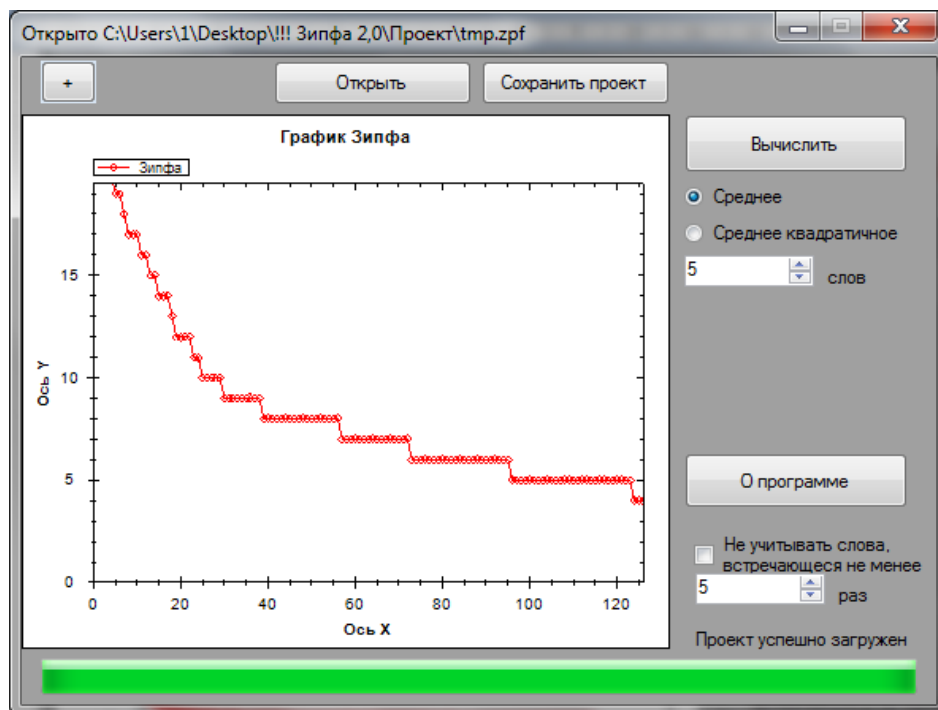


Рис. 4. Скорректированный график гиперболы

Результирующие слова Зипфа для расчёта среднего показаны на рисунке 5, а среднеквадратичного – на рисунке 6.

Файл Правка Формат Вид		
Справка		
вывод	1	
образом	1	
инстанций		1
состава	1	
наличии	1	
таким	1	
зале	1	
cache	1	
нарушением		1
datatype		1

Рис. 5. Слова Зипфа при расчёте среднего

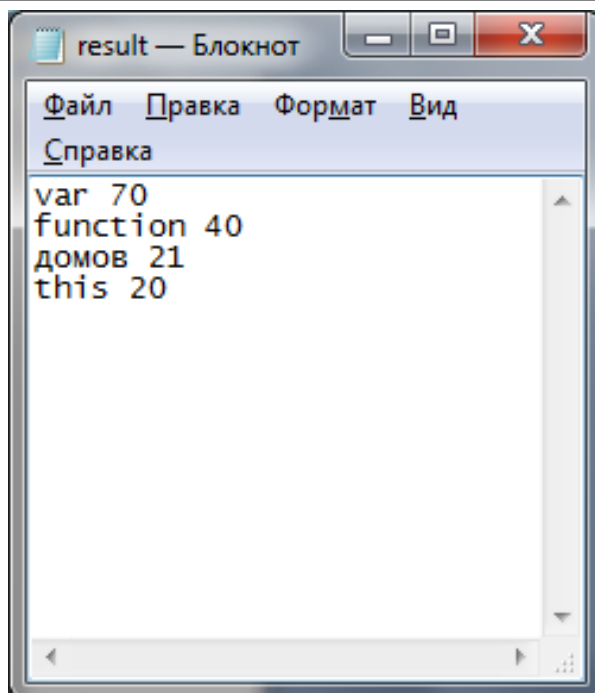


Рис. 6. Слова Зипфа при расчёте среднеквадратичного значения

При анализе текстов на основе второго закона Зипфа необходимо исключить так называемые «стоп-слова» (предлоги, союзы и т.п.). Для этого в программе предусмотрена возможность ведения словаря исключений (рис. 7).

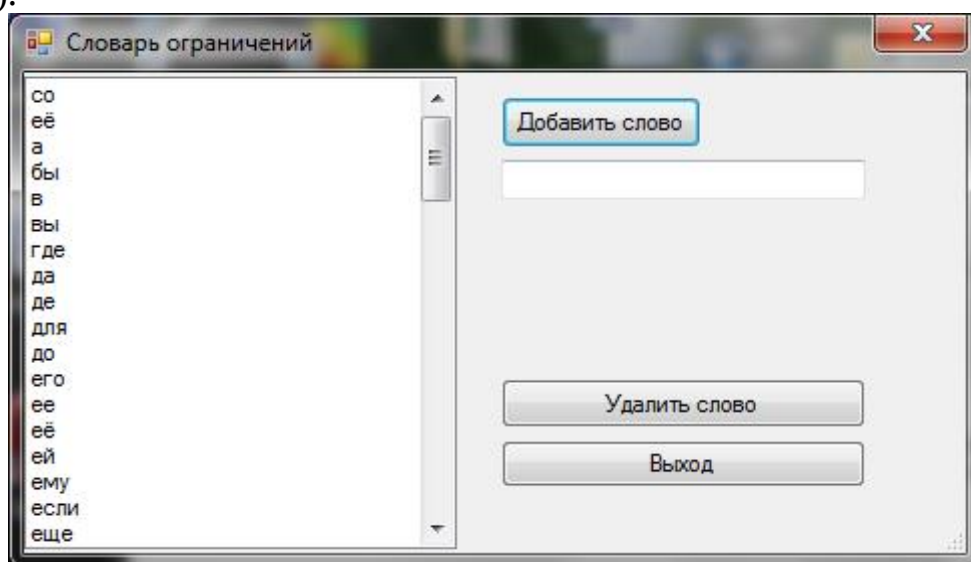


Рис. 7. Окно редактирования словаря исключений

Варьируя интервалы анализа, количество значимых и незначимых слов, а также выбирая режим расчётов, мы имеем возможность провести подробный анализ любого произвольного текста с целью выявления его внутренней смысловой структуры [5].

Список литературы

1. Хорошилов А.В., Селетков С.Н. Мировые информационные ресурсы // СПб.: ПИТЕР, 2004. – 366 с.
2. Кохонен Т. Самоорганизующиеся карты: пер. 3-го англ. изд. – М.: БИНОМ. Лаборатория знаний, 2008. – 655 с.: ил. – (Адаптивные и интеллектуальные системы).
3. Мешкова Е.В. Методика построения классификатора текста на основе гибридной нейросетевой модели // Известия ЮФУ. Технические науки. Тематический выпуск «Интеллектуальные САПР». – Таганрог: Изд-во ТТИ ЮФУ. – 2008. – № 4 (81). – 212–216 с.
4. Мешков В.Е., Мешкова Е.В., Светличный Р.А. Интеллектуальная поисковая система // Теория, методы и средства измерений, контроля и диагностики: материалы II Междунар. научн.-практич. конф., г. Новочеркасск, 21 сентября 2001 г.: в 4 ч. / Южно-Рос. гос. ун-т (НПИ). – Новочеркасск: Темп, 2001. – Ч. 4. – 76 с.
5. Мешков В.Е., Мешкова Е.В. Построение гибридной модели на основе семантической и ассоциативной сетевых парадигм // Радиоэлектроника, телекоммуникации, информационные технологии: межвуз. сб. науч. трудов; Южно-Рос. гос. ун-т экономики и сервиса / ред. кол.: Н.Н. Прокopenко [и др.]. – Шахты: Изд-во ЮРГУЭС, 2004. – 30–33 с.

УДК 004.01

В.Е. Мешков, В.С. Чураков

ПРОГРАММА ИССЛЕДОВАНИЙ В ОБЛАСТИ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И КОГНИТОЛОГИИ В РАМКАХ СЕМИМЕРНОЙ ПАРАДИГМЫ А.В. КОРОТКОВА

В статье предлагается программа исследований на базе семимерной парадигмы А.В. Короткова. Показывается, что могут быть получены значимые результаты в области информатики, когнитологии, искусственного интеллекта.

Ключевые слова: семимерная парадигма, когнитология, информатика, алгебры логики.

Предлагаемая нами «Программа исследований» базируется на семимерной парадигме А.В. Короткова.

Анатолий Васильевич Коротков проблемой семимерного пространства (собственно евклидового и псевдоевклидового) занимается с 1988 г. В 2001 г. А.В. Коротков защитил докторскую диссертацию «Элементы се-

мимерного векторного исчисления в задачах математического моделирования физических и технических объектов». Специальности: 05.13.01 – «Системный анализ, управление и обработка информации (по отраслям)», 05.13.18 – «Математическое моделирование, численные методы и комплексы». Диссертационный совет рекомендовал Высшей Аттестационной Комиссии Российской Федерации включить в учебные программы основные сведения о семимерном векторном исчислении для студентов механико-математических и физико-математических специальностей вузов.

Семимерная парадигма А.В. Короткова представляет собой семимерное пространство (собственно евклидово и псевдоевклидово). Оно обусловлено тем обстоятельством, что математики Новосибирской школы показали, что трёхмерная алгебра является подалгеброй только семимерной алгебры. Нужно было рассматривать семимерный вариант со скалярным и векторным произведением двух векторов, т.е. семимерную векторную алгебру – в философском отношении «истинной середины» (следуя Мих. Лифшицу) – в отношении множества многомерных концепций пространства. (Кроме того, кватернионы организуют координацию векторизованных явлений в трёхмерном пространстве, в котором существует лишь семь различных систем координат.) Это – исходная *эпистемологическая парадигма, которой мы будем следовать.*

На основе результатов, полученных в многолетних исследованиях по семимерной парадигме, в последние годы А.В. Коротков опубликовал работы, позволившие применить семимерную парадигму к информационным технологиям, искусственному интеллекту и когнитологии (в рамках семимерной парадигмы А.В. Коротков разработал: дискретные алгебры [многомерные целочисленные алгебры], многозначные алгебры логики и небулевы алгебры логики [2–4]), способные ещё более понизить конкурентную способность гипотетических квантовых компьютеров по отношению к универсальным электронным вычислительным машинам.

Разработки А.В. Короткова позволяют наполнить новым содержанием, прежде всего, два традиционных подхода: логический и структурный.

Логический подход. Основой для логического подхода служит Булева алгебра. Она появилась ещё в XIX в. (Своё дальнейшее развитие Булева алгебра получила в виде исчисления предикатов, в котором она расширена за счёт введения предметных символов, отношений между ними, кванторов существования и всеобщности). В вычислительной технике она применяется с середины XX в. Булева алгебра двузначна и одномерна. Это числа с двумя состояниями, которые условно называют нуль и единица. Они дают соответствующие операции сложения и умножения этого числа, в результате булева алгебра обладает целым рядом полезных, очень важных свойств, позволивших широко применять её на практике, в алгебре логики, а также (самое главное) – в технике логических, арифметических

и преобразовательных устройств. Булева алгебра хорошо разработана и изложена, говорить о ней много не надо (но попутно стоит отметить такое сравнительно новое направление, как нечёткая логика). Необходимо отметить прецедент, который возникает в булевой алгебре.

Во-первых, наличие операций сложения не сопровождается операцией вычитания, то есть отсутствует противоположность операции сложения. Булевому числу нельзя сопоставить противоположное число. В алгебре, например, действительных чисел всё обстоит иначе. Эта алгебра характеризует поле – математическое понятие, набор математических операций, одна из которых – операция вычитания. Так вот, имеется система в теории сравнений, которая может работать с классом вычетов по модулю. Эта теория хорошо разработана и изложена в литературе. Она имеет возможность построения чисел по модулю 2 классов сравнений и классов вычетов по модулю два. Это – та же система с двумя числами – нуль и единица, но эта система не имеет уже операций вычитания и в результате имеет отличающуюся от булевой алгебры операцию сложения, где единица плюс единица в этой алгебре есть нуль, в то время как в булевой алгебре единица плюс единица есть единица.

Это очень существенное отличие, позволяющее построить новую алгебру. Эта алгебра обладает целым рядом полезных свойств, теми же, что и, например, в алгебре действительных чисел, хотя она дополнена свойствами чисто логических систем. В этой алгебре A равняется нулю, а не $\bar{A}$, как в булевой алгебре, если A – число. Отличительные свойства этой небулевой одномерной алгебры от булевой одномерной алгебры характеризует целый ряд возможностей и создаёт целый ряд алгебр. Это в отношении одномерной алгебры булевой и небулевой.

В принципе, булева алгебра может быть расширена до многомерного варианта, до N -мерного, где N – произвольное число, путём применения операции умножения. Умножение в этой алгебре прямое – умножение двух чисел. В одной из публикаций показано, что булева алгебра может быть N -мерна, то есть число может быть записано в N -мерной форме. Операнды в N -мерной форме есть результаты операций в N -мерной форме. Это создаёт возможности использования этой алгебры по ряду назначений, в частности при построении логических многомерных устройств либо логически-арифметических многомерных устройств. Однако прямое произведение двух величин всё-таки обладает некоторыми существенными недостатками, поэтому в практике действительных чисел используют не только прямое произведение двух величин, но и произведение многомерных величин, построенных не по способу прямого произведения. Такими числами, кроме действительных одномерных чисел, являются комплексные числа, например двумерные числа, кватернионные четырёхмерные числа, октанионы восьмимерные числа, а также числа, характеризующие векторные

алгебры, одномерные векторные алгебры, трёхмерные векторные алгебры, семимерные векторные алгебры. То есть в алгебре действительных чисел имеется целый ряд возможностей расширения, но поскольку система сравнений классов вычетов по модулю обладает свойствами, близкими к свойствам действительных чисел, в частности, обладает вычитанием, можно строить алгебры логики многомерной, используя те же процедуры для произведения чисел, что и алгебры многомерной для действительных чисел. В частности, процедура удвоения Гамильтона может быть использована для построения двумерной алгебры логики и одномерной векторной алгебры. Та же процедура удвоения Гамильтона, применённая к комплексным, логическим числам, даст четырёхмерные логические числа и трёхмерную векторную алгебру. Та же процедура удвоения Гамильтона, применённая к кватернионным числам, даст октанионную систему чисел – логических чисел и семимерную векторную алгебру логики, то есть известная процедура умножения по отношению к числам классов вычетов по модулю даёт возможность построить целый ряд совершенно новых алгебр. Следовало бы отметить ещё одно важное применение – речь идёт о дискретных алгебрах с бесконечным модулем в данном случае, вернее, это – алгебра действительных чисел, но только расширена она до четырёх восьмимерных (одно-, трёх- и семимерный вариант). Дело в том, что в векторных алгебрах, алгебрах кватернионов, комплексных чисел и октанионов используются операции сложения и умножения, а также операции скалярного произведения и векторного произведения двух векторов векторной алгебры. Если числа занимают чисто дискретный ряд значений, например приобретают только целые значения, то результаты всех практических операций будут принимать целые значения. То есть эта алгебра в значительной степени воспроизводит дискретную алгебру, если речь не идёт об извлечении квадратного корня. Скалярное произведение, векторное произведение, смешанное произведение трёх векторов, двойное векторное произведение трёх векторов, есть и другие многомерные операции, которые будут иметь целочисленные значения, многомерные целочисленные значения, это может иметь существенное применение для описания дискретных величин, которые в последнее время на протяжении уже более ста лет широко используются физиками.

Структурный подход – это не что иное, как попытки построения ИИ путём моделирования структуры человеческого мозга. Одной из первых таких попыток был перцептрон Френка Розенблатта. Основной моделируемой структурной единицей в перцептронах (как и в большинстве других вариантов моделирования мозга) является нейрон. Позднее возникли и другие коннекционистские модели, которые большинству известны под термином нейронные сети (НС) и их реализации – нейрокомпьютеры. Эти модели различаются по строению отдельных нейронов, по топологии

связей между ними и по алгоритмам обучения. Среди наиболее известных сейчас вариантов НС можно назвать НС с обратным распространением ошибки, сети Кохонена, сети Хопфилда, стохастические нейронные сети. В более широком смысле такой подход известен как коннективизм. Различия между логическим и структурным подходом не столь принципиальны, как это может показаться на первый взгляд. Алгоритмы упрощения и вербализации нейронных сетей преобразуют модели структурного подхода в явные логические модели. С другой стороны, ещё в 1943 г. Маккалок и Питс показали, что нейронная сеть может реализовывать любую функцию алгебры логики [16].

На основании вышеизложенного сформулируем программу исследований в области информационных технологий, искусственного интеллекта и когнитологии в рамках семимерной парадигмы А.В. Короткова.

В области *информационных технологий* в рамках семимерной парадигмы возможно решение следующих задач:

- разработка математических основ для создания многомерного булевого и небулевого логического базиса элементов и устройств вычислительной техники (включая нейрочипы, нейрокомпьютеры и нейросети);
- разработка принципиальных схем для реализации полного набора базовых элементов вычислительных устройств на основе многомерного булевого и небулевого логического базиса;
- разработка новых алгоритмов и программного обеспечения на основе многомерных булевых и небулевых логик;
- в робототехнике возможно построение систем представления и анализа пространственных объектов в семимерном пространстве координат (трёхмерные пространственные координаты X, Y, Z , цветовое восприятие, обоняние, осязание, аудиоинформация).

В области *искусственного интеллекта* в рамках семимерной парадигмы возможно решение следующих задач:

- разработка теоретических основ построения многомерных булевых и небулевых формальных логик;
- построение экспертных систем на базе многозначных и многомерных булевых и небулевых формальных логик;
- применение семимерной парадигмы при построении многовекторных систем управления базами данных (МСУБД);
- применение семимерной парадигмы для реализации поисковых систем с многовекторным индексированием гипертекста;
- применение семимерной парадигмы для формирования производственных моделей предметных областей в интеллектуальных информационных системах;

- открывается возможность для следующих вариантов построения нейросетей [13]:
 - а) на небулевых алгебрах логики;
 - б) на многозначных алгебрах логики;
 - в) на многомерных алгебрах логики.

Работы в этих направлениях уже начаты (например [7–10]).

Методика построения ИИ на трёхмерных логиках (*стандартный ИИ*) и построение *ИИ высокого уровня* – на алгебраических логиках более высокого порядка должна быть социальной, примерно так, как это описано в романе Й. Макдональда «Река Богов» [5], и многомерной.

Поскольку представления о природе человеческого сознания, полученные в когнитивной психологии (и проецируемые на модели искусственного интеллекта), по замечанию И.З. Цехмистро, «не идут дальше выяснения функциональных сторон его деятельности: памяти, логико-вычислительных операций, способности к прогнозированию и т.п., которые с той или иной степенью достоверности могут быть смоделированы в различных кибернетических устройствах» [17, с. 4], то в области *когнитологии (когнитивного моделирования)* в рамках семимерной парадигмы возможно:

- построение многомерной модели сознания на основе многомерных булевой и небулевой алгебр А.В. Короткова [6]. (Открывается возможность построения [многомерной] *топологии сознания*, для чего лучше всего подходят работы: Л. Литвака «Жизнь после смерти»: предсмертные переживания и природа психоза. Опыт самонаблюдения и психоневрологического исследования» [11] и В. Никитаева «Герменевтика смерти» [12]);
- появляется возможность на основе пифагоровых чисел моделировать представление сознания и мышления, а также перейти от евклидовых представлений (моделирования работы мозга) – к неевклидовым, дополнительным к евклидовым;
- по-новому можно подойти к проблеме формализации знаний;
- кроме того, данный подход можно применить к *теории автоматов, в нечётком моделировании и управлении* [14] и использовать *методы [небулевой, многозначной и многомерной] алгебры логики в математической физике*.

Предложенная выше исследовательская программа представляет собой, по сути, новое научное направление на многомерной математической базе, которое можно применить и в альтернативном подходе [1; 15].

Список литературы

1. Клименко А.В. Основы естественного интеллекта. Рекуррентная теория самоорганизации. Версия 3. – Ростов н/Д.: Изд-во Ростовского университета, 1994. – 304 с.
2. Коротков А.В., Чураков В.С. Многомерные булевы алгебры // Теоретико-философские аспекты трёхмерного и семимерного пространств (собственно евклидова и псевдоевклидова). – Новочеркасск: Набла, 2007. – 194 с. – (С. 180–185).
3. Коротков А.В. Многозначные алгебры логики // Информационные системы и технологии. Теория и практика: сб. научн. тр. – Шахты: Изд-во ЮРГУЭС, 2008. – С. 17–23.
4. Коротков А.В. Небулевы алгебры логики / / Информационные системы и технологии. Теория и практика: сб. научн. тр. – Шахты: Изд-во ЮРГУЭС, 2008. – С. 23–29.
5. Макдональд Й. Река Богов / пер. с англ. С. Минкина. – М.: АСТ: АСТ МОСКВА: Транзиткнига, 2006. – 669 с.
6. Мешков В.Е., Чураков В.С. Многомерная модель сознания на основе многомерных булевой и небулевой алгебр А.В. Короткова // «Интеллектуальные системы» (AIS'08) и «Интеллектуальные САПР» (CAD-2008): тр. Междунар. науч.-технич. конф.: в 4 т. – М.: Физматлит, 2008. – Т. 2. – С. 159–161.
7. Мешков В.Е., Чураков В.С. Многомерные алгебры А.В. Короткова для нейросетей и нейрокомпьютеров // AIS-IT'09: тр. Конгресса по интеллектуальным системам и информационным технологиям: науч. изд-е: в 4 томах. – М.: Физматлит, 2009. – Т. 2. – 568 с. – (С. 112–114).
8. Мешков В.Е., Чураков В.С. Пифагоровы числа в нейросетях // Труды Конгресса по интеллектуальным системам и информационным технологиям AIS-IT'09: в 4 т. – М.: Физматлит, 2009. – Т. 2. – 568 с. – (С. 115–119).
9. Многозначные логики и их применения. Т. 1. Логические исчисления, алгебры и функциональные свойства / сост. О.М. Аншаков, Д.В. Виноградов, В.К. Финн; под ред. В.К. Финна. – М.: Изд-во ЛКИ, 2008. – 504 с.
10. Многозначные логики и их применения. Т. 2. Логики в системах искусственного интеллекта / сост. О.М. Аншаков, Д.В. Виноградов, В.К. Финн; под ред. В.К. Финна. – М.: Издательство ЛКИ, 2008. – 240 с.
11. Литвак Л. «Жизнь после смерти»: предсмертные переживания и природа психоза. Опыт самонаблюдения и психоневрологического исследования / под ред. и со вступит. статьёй Д.И. Дубровского. – Изд. 2-е, перераб. и доп. – М.: Канон+, РООИ «Реабилитация», 2007. – 672 с.

12. Никитаев В. Герменевтика смерти // Логос. – 2005. – № 2 (47). – С. 193–211.
13. Осовский С. Нейронные сети для обработки информации / пер. с польск. И.Д. Руденко. – М.: Финансы и статистика, 2004. – 344 с.: ил.
14. Пегат А. Нечёткое моделирование и управление: пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2009. – 798 с.: ил.
15. Поликарпов В.С. Философские проблемы искусственного интеллекта / В.С. Поликарпов, В.М. Курейчик, Е.В. Поликарпова [и др.]. – М.: Физматлит, 2008. – 266 с.
16. Солсо Р. Когнитивная психология. – 6-е изд. – СПб.: Питер, 2006. – 589 с.: ил.
17. Цехмистро И.З. Поиски квантовой концепции физических оснований сознания. – Харьков: Вища школа: Изд-во при Харьк. ун-те, 1981. – 176 с.

РАЗДЕЛ 3

ОПТОИНФОРМАТИКА

УДК 004.387; 535.417

З.С. Бекашева, В.Н. Васильев, А.А. Востриков, А.В. Павлов

СТАТИСТИЧЕСКИЕ СВОЙСТВА ГИПОТЕЗ, ПОРОЖДАЕМЫХ ДВУСЛОЙНОЙ НЕЙРОННОЙ СЕТЬЮ С МАТРИЦЕЙ ДВУНАПРАВЛЕННЫХ СВЯЗЕЙ, РЕАЛИЗУЕМОЙ МЕТОДОМ ГОЛОГРАФИИ ФУРЬЕ*

Применительно к выдвижению гипотез двуслойной нейронной сетью с матрицей связей, реализуемых методом голографии Фурье, исследовано влияние фильтрации, возникающей вследствие ограниченности динамического диапазона голографической регистрирующей среды, на статистические характеристики паттерна внутренней репрезентации гипотезы. При численном моделировании использованы модели фильтров, адекватных реальным характеристикам голографических регистрирующих сред. Получены зависимости, связывающие оценку стационарности с характеристиками процесса и фильтра.

Ключевые слова: оптические нейронные сети, голография, творческое мышление, предсказание.

Введение. Одна из актуальных задач искусственного интеллекта – реализация механизмов творческого мышления. По мнению многих исследователей, творческие способности носителя интеллекта зависят в значительной степени от двух факторов – развитости образного (правополушарного) мышления [1–8] и способности к «погружению в хаос», т.е. переходу к хаотической динамике нейронной активности, и выходу из него [9–11].

Один из возможных подходов к реализации образного мышления основан на том, что мозг как нейронная сеть (НС) обрабатывает картины нейронной активности, формирующиеся в коре мозга при восприятии сенсорами информации. Эти картины нейронной активности суть паттерны

* Работа выполнена при поддержке РФФИ, гранты 09-01-00165-а и 09-02-00223-а.

внутренней репрезентации воспринимаемой информации (ПВР). Соответственно, в рамках концепции образного мышления механизм решения задач мозгом может рассматриваться следующим образом:

- условия встреченной задачи представляются нейронной сети в виде ПВР условий задачи (воспринимаемой информации);
- ответ формируется нейронной сетью как ПВР ответа – результат обработки нейронной сетью ПВР условий задачи на основе имеющихся знаний, представленных в виде матрицы межнейронных связей.

Согласно результатам, полученным в когнитивной науке, жизнедеятельность организма направлена на восстановление субъективной целостности внутренней картины мира, нарушаемой воспринимаемой извне информацией [10]. В рамках такой концепции воспринимаемая информация является субъективно проблемной, т.е. ПВР такой информации воспринимается как ПВР задачи, которую надо решить, если он не соответствует имеющимся знаниям – воспринимается как искажённый или дефектный вариант известного образа. Соответственно, процесс решения задачи – достройка или восстановление субъективно целостного образа, а ответ – достроенный и исправленный ПВР.

В рамках такого подхода творческая задача отличается от задачи нетворческой тем, что ПВР ответа не может быть сформирован НС посредством обработки ПВР условий задачи на основе имеющихся знаний, представленных в виде матрицы связей, методами какого-либо из формальных выводов (дедуктивного, абдуктивного или индуктивного) [12]. Поэтому один из возможных методов решения творческой задачи предполагает генерацию большого числа образов, незначительно отличающихся друг от друга, но принадлежащих одной области фазового пространства НС [7, 9]. Эти образы, не связанные с ПВР условий и знаниями правил вывода, суть творческие гипотезы – варианты ответов на встреченную задачу. Решением творческой задачи может быть одна из порождённых гипотез, удовлетворяющая критерию адекватности реальности [13]. «Погружение в хаос», т.е. переход к хаотической динамике нейронной активности, как раз и является методом генерации большого числа новых картин нейронной активности, принадлежащих одной области фазового пространства НС – странному аттрактору.

Основываясь на этих посылах, в работах [13–15] был предложен подход к реализации механизма решения творческих задач двуслойной НС с матрицей двунаправленных связей, формируемой методом голографии Фурье (рис. 1).

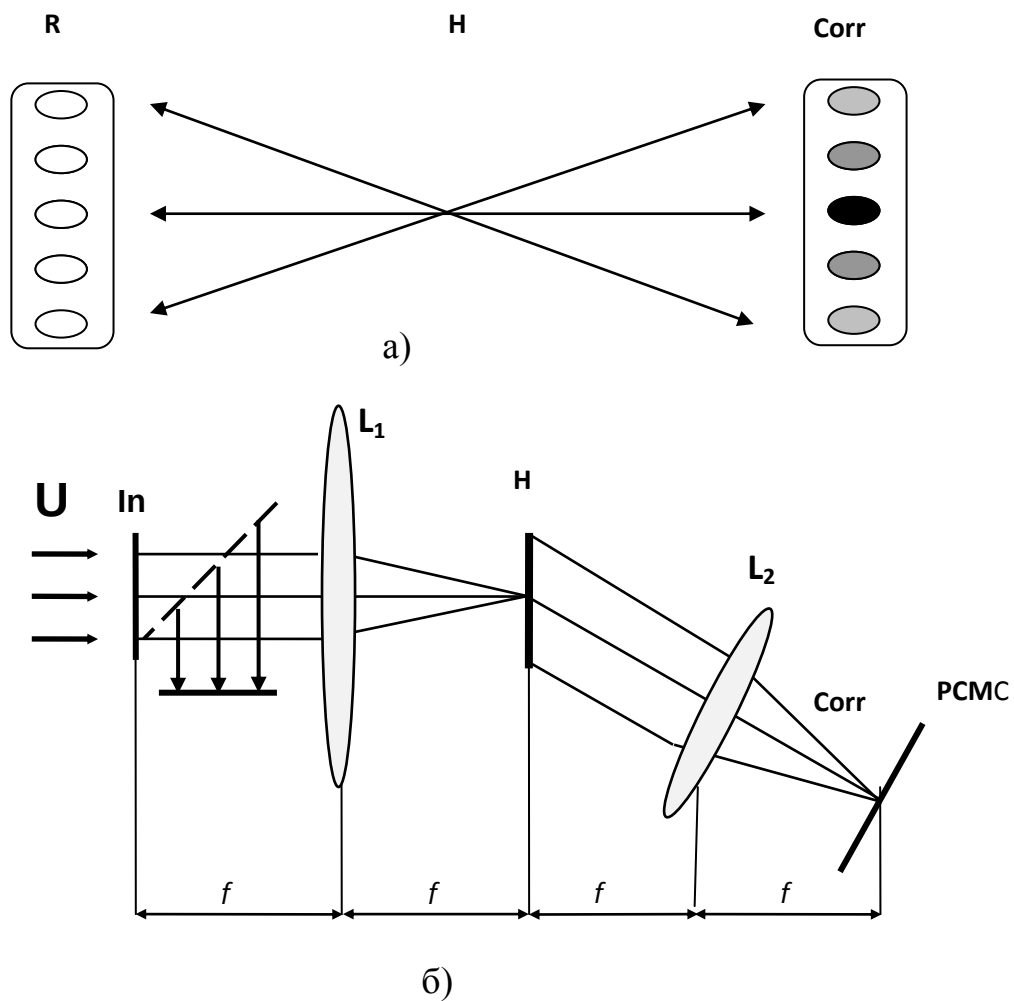


Рис. 1. Схема двуслойной нейронной сети с матрицей двунаправленных связей (а) и её реализация методом голографии Фурье (б):

R – слой репрезентаций, соответствующий входной плоскости In ;

$Corr$ – слой корреляций (корреляционная плоскость);

L_1 и L_2 – Фурье-преобразующие линзы с фокусными расстояниями f ;

H – матрица межнейронных связей (голограмма); РСМ – фазосопрягающие зеркала во входной и корреляционной плоскостях

Особенности этого подхода следующие:

1. ПВР воспринимаемой информации $Im_{Per}(x,y)$ (условия задачи) формируется в слое репрезентаций R , знания реализуются в виде матрицы связей нейронных слоёв H (голограммы Фурье), а в слое корреляций $Corr$ формируется функция взаимной корреляции условий задачи и имеющихся знаний.

2. Мерой «проблемности» воспринимаемой информации служит мера сходства ПВР $Im_{Per}(x,y)$ с имеющимися знаниями, в качестве которой выступает амплитуда глобального максимума автокорреляционной функции (ГМ АКФ). Если амплитуда ГМ АКФ превышает некоторый порог, то

$\text{ImPer}(x, y)$ воспринимается как знакомый и субъективно целый $\text{ImR}(x, y)$ – проблемы, т.е. задачи, которую надо решить для того, чтобы восстановить целый ПВР, нет.

3. Если функция взаимной корреляции ПВР условий задачи и знаний имеет выраженный ГМ АКФ, но его амплитуда меньше порога, то $\text{ImPer}(x, y)$ знаком, но субъективно неполон или искажён. Возникает задача – достроить или исправить $\text{ImPer}(x, y)$ до эталонного $\text{ImR}(x, y)$. В этом случае условия задачи соответствуют имеющимся знаниям, и она может быть решена на основе имеющихся знаний либо вспоминанием известного ответа (автоассоциативная память), либо его построением методом регрессии условий задачи по знаниям [16].

4. Если ПВР условий задачи не соответствует имеющимся знаниям, то ГМ АКФ либо отсутствует, либо его амплитуда не превосходит амплитуды боковых максимумов. Ни вспомнить правильный ПВР, ни вывести его не удаётся. ПВР ответа можно только придумать (увидеть во сне, например) – это творческая задача. Для того чтобы «придумать» ПВР (ответ), НС переходит в режим хаотической активности, в котором в слое R на каждой итерации $R \rightarrow \text{Согг} \rightarrow R$ генерируется новый образ – творческая гипотеза. Всё множество генерируемых гипотез принадлежит одному странному аттрактору, определяемому как характеристиками НС, так и имеющимися знаниями. Зависимость характеристик странного аттрактора от характеристик НС и знаний может служить механизмом страхования от порождения химер, т.е. образов, противоречащих фундаментальным законам мироздания.

5. Следующий этап – верификация порождённых гипотез по критерию адекватности реальности для выбора одной гипотезы – решения задачи.

Свойства ПВР гипотез, порождаемых в режиме хаотической активности, определяются тем, что в НС (рис. 1) матрица связей нейронных слоёв R и Согг , хранящая знания, формируется как голограмма Фурье:

$$H_{R \rightarrow C}(\nu_x, \nu_y) = F^*(\text{Im}_R(x, y)), \quad (1)$$

где (ν_x, ν_y) – пространственные частоты; астериск $*$ обозначает комплексное сопряжение; F – преобразование Фурье.

При формировании в слое R ПВР $\text{ImPer}(x, y)$ в плоскости Согг в результате прохождения возбуждения (света) $R \rightarrow H \rightarrow \text{Согг}$ формируется функция взаимной корреляции условий задачи $\text{ImPer}(x, y)$ и знаний $\text{Im}_R(x, y)$:

$$C(\xi, \eta) = \text{Im}_{\text{Per}}(x, y) \otimes \text{Im}_R(x, y) = F(F(\text{Im}_{\text{Per}}(x, y))F^*(\text{Im}_R(x, y))),$$

где ξ, η – координаты в плоскости корреляций. Для того чтобы не потерять разрешение, на этапе обратного прохождения света от плоскости Согг к плоскости R должна использоваться голограмма с инверсной зависимостью локальной дифракционной эффективности от частоты (метод де-свёртки):

$$H^{-1}(\nu_x, \nu_y) = \left(\frac{1}{F^*(\text{Im}_R(x, y))} \right). \quad (2)$$

Тогда в предположении линейности активационных функций C -нейронов ореол вокруг ПВР $\text{Im}_R(x, y)$, восстановленного в слое R в результате одной итерации $R \rightarrow H \rightarrow \text{Corr} \rightarrow H^{-1} \rightarrow R$, описывается выражением:

$$\text{Im}(x_k, y_l) = F \left(\frac{F(C(x_k + \xi, y_l + \eta))}{F^*(\text{Im}_R(x, y))} \right), \quad (3)$$

где (x_k, y_l) – координаты в плоскости R вне области определения $\text{Im}_R(x, y)$ (ореол). В работах [16, 17] показано, что выражение (3) посредством двойного применения преобразования Фурье и теоремы Бореля о свёртке может быть получено из известной в теории случайных процессов модели множественной линейной регрессии [18–20], представленной в виде:

$$\text{Im}(x_k, y_l) = \int_{x_{\text{Min}}}^{x_0} \int_{y_{\text{Min}}}^{y_0} \text{Im}(x_0 - x, y_0 - y) a(x, y) dx dy,$$

где весовая функция $a(x, y)$ находится из решения интегрального уравнения:

$$\int_{x_{\text{Min}}}^{x_0} \int_{y_{\text{Min}}}^{y_0} a(x, y) C(|x - \xi|, |y - \eta|) dx dy = C(x_k + \xi, y_l + \eta).$$

Если изображение $\text{Im}(x, y)$ описывается как реализация однородного случайного поля, то (3) описывает наилучшую по критерию минимума среднего квадрата ошибки оценку.

Для перехода к режиму хаотической активности по сценарию Фейгенбаума активационные функции C -нейронов должны иметь нелинейность соответствующего вида, в этом случае выражение (3) переписывается в виде:

$$\text{Im}_N(x_k, y_l) = F \left(\frac{F(Nl(C_{N-1}(x_k + \xi, y_l + \eta)))}{F^*(\text{Im}_R(x, y))} \right), \quad (4)$$

где Nl – нелинейные активационные функции C -нейронов, а N – номер итерации.

Известно, что динамический диапазон регистрирующих сред, могущих быть использованными для записи голограмм (1) и (2), всегда ограничен, а пространственно-частотные спектры реальных изображений имеют перепад амплитуд, существенно превышающий динамический диапазон регистрирующих сред. В результате на голограмме записывается только ограниченный диапазон амплитуд и, соответственно, ограниченный интервал частот, выбор которого каждый раз определяется задачей. Соответственно, и гипотезы (3) и (4) строятся не для исходного $\text{ImPer}(x, y)$, но для профильтрованного голограммой ПВР. Например, на приведённых в [17] экспериментально полученных изображениях хорошо виден результат узкополосной фильтрации – подчёркивание контуров.

Кроме того, физические методы воплощения инверсного фильтра предполагают не непосредственную реализацию выражения (2), а запись согласованного фильтра и добавление к нему амплитудного фильтра:

$$H^{inv}(\nu_x, \nu_y) = \left(\frac{1}{(F(\text{Im}_R(x, y))F^*(\text{Im}_R(x, y)))} F^*(\text{Im}_R(x, y)) \right), \quad (5)$$

что также имеет следствием сужение полосы частот.

Эта неизбежная пространственно-частотная фильтрация влияет также и на такие характеристики порождаемых ПВР гипотез, существенные с точки зрения модели, как матожидание и однородность (стационарность). Изменение статистических свойств гипотезы по сравнению со свойствами ПВР условий задачи следует рассматривать с точки зрения адекватности гипотезы условиям задачи. Поэтому представляет интерес оценка влияния частотной фильтрации на статистические характеристики изображения применительно к свойствам реальных голографических сред, используемых для реализации матрицы межнейронных связей.

В настоящей статье рассмотрено влияние частотной фильтрации на статистические характеристики порождаемых творческих гипотез применительно к двум типам голограмм, условно называемых низкочастотными и высокочастотными. В дальнейшем изложении, приняв для упрощения выкладок, но без потери общности допущение о разделимости переменных в функциях $\text{Im}(x, y)$, описывающих ПВР, будем по преимуществу использовать термины «процесс» и «стационарность», применимые к сечениям $\text{Im}(x)$.

Классификация голограмм

Не вдаваясь в детали, далёкие от интересов большинства читателей журнала, отметим вкратце, что передаточной функцией $4f$ схемы голографии Фурье (рис. 1) является зависимость локальной дифракционной эффективности голограммы Фурье от частоты $H(\nu_x, \nu_y)$, которая определяется условиями записи голограммы, в первую очередь частотой равенства локальных амплитуд сигнального и опорного пучков ν_0 [21–22]. В первом приближении выделим два типа голограмм, которые условно назовём «высокочастотными» и «низкочастотными». Если на частотах ниже ν_0 наблюдается инверсная зависимость $H(\nu_x, \nu_y)$ от амплитуды пространственно-частотного спектра эталона $F^*(\text{Im}(x, y))$, т.н. «провал» дифракционной эффективности, то голограмму будем называть высокочастотной. Если же частота ν_0 достаточно низка или близка к нулевой, то «провал» дифракционной эффективности отсутствует – такую голограмму будем называть низкочастотной. Рассмотрим влияние высокочастотной и низкочастотной фильтрации по отдельности.

Низкочастотная фильтрация

Низкочастотные голограммы представляют интерес при наличии высокочастотного (или «белого») шума, поскольку позволяют существенно увеличить отношение сигнал/шум. При такой фильтрации в первую очередь актуально изменение стационарности процесса по дисперсии. Оценку стационарности по дисперсии будем проводить на основе сравнения корреляционных функций $Cor1_j$ и $Cor2_j$ двух фрагментов реализации с учётом нормировки по дисперсии D_1 и D_2 по формуле (для упрощения выражения рассмотрен одномерный случай):

$$V_D = \frac{1}{N-1} \cdot \sum_0^{N-1} (Cor1_j - Cor2_j)^2 \cdot \left(\frac{\max(D_1, D_2)}{\min(D_1, D_2)} \right)^2. \quad (6)$$

Методом численного моделирования было проведено сравнение оценок стационарности для четырёх типов пространственно-частотных фильтров – скользящего среднего, прямоугольного, экспоненциального и описываемого функцией Гаусса. Прямоугольный фильтр выбран из-за простоты своей реализации, а экспоненциальный и гауссообразный – в силу адекватности реальным передаточным функциям [21]. Описание фильтров приведено в таблице 1, значения параметров, входящих в аналитические выражения, подбирались так, чтобы при одном значении параметра θ значения всех функций были равны 0,25. Использовались реализации случайного процесса с экспоненциальным спектром и явно заданной нестационарностью по дисперсии.

Таблица 1

Низкочастотные фильтры

Скользящий средний	Прямоугольный	Экспоненциальный	Гауссообразный
$f_m(v) = \text{Sin c}\left(\frac{\pi v \theta}{210}\right)$	$f_R(v) = \text{Rect}(v_\theta)$	$f_e(v) = \exp\left(-\frac{v \theta}{120}\right)$	$f_G(v) = \exp\left(-\frac{(v \theta)^2}{19880}\right)$

На рисунке 2 приведены полученные зависимости оценки стационарности от значения параметра θ , определяющего ширину фильтра по уровню 0,25, для четырёх видов фильтров при значении дисперсии $D=3$.

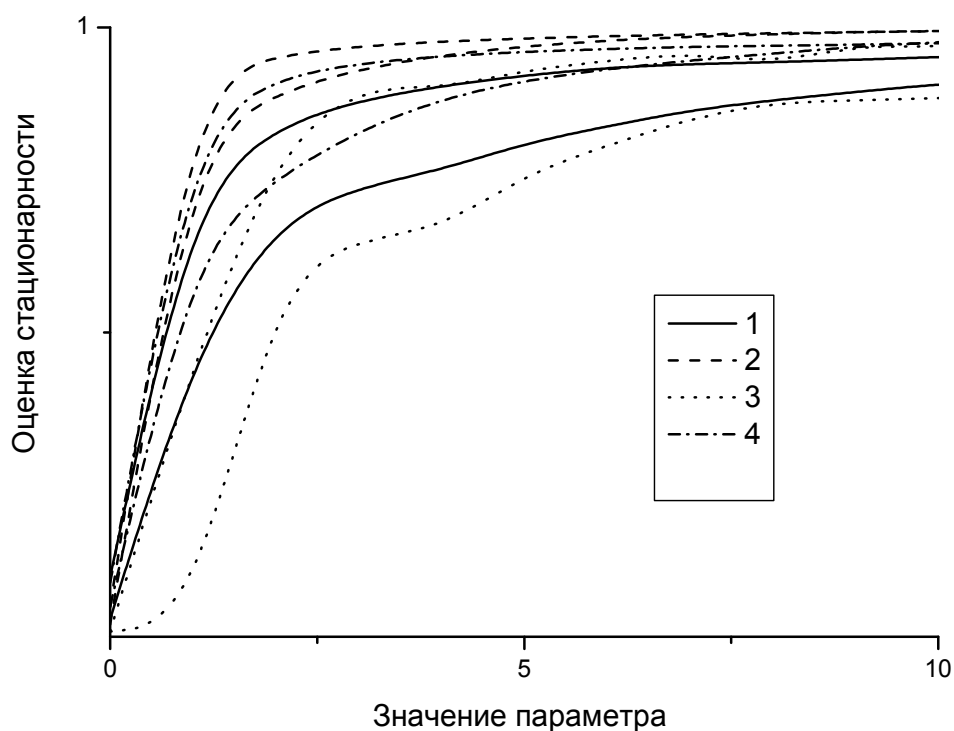


Рис. 2. Зависимости оценок стационарности по дисперсии от характеристик низкочастотных фильтров при исходной дисперсии процесса $D = 3$: 1 – фильтр скользящего среднего; 2 – экспоненциальный фильтр; 3 – прямоугольный фильтр; 4 – Гауссов фильтр

При увеличении дисперсии исходного процесса рост оценок стационарности происходил более плавно – кривые на графике опускались вниз. Наибольший рост оценок стационарности по дисперсии дал самый узкий, т.е. экспоненциальный, фильтр, следующий – гауссообразный.

Отметим, что при возрастании параметра θ уменьшается и дисперсия, т.е. имеет место «сглаживание» процесса, он стремится к своему матожиданию.

Влияние высокочастотной и полосовой фильтрации на оценку стационарности по матожиданию

При численном моделировании полосовой фильтрации использованы случайные процессы с приведёнными на рисунке 3 двумя видами спектров, адекватно аппроксимирующих спектры значительной части реальных изображений и сигналов – экспоненциальным и экспоненциальным с локальным максимумом.

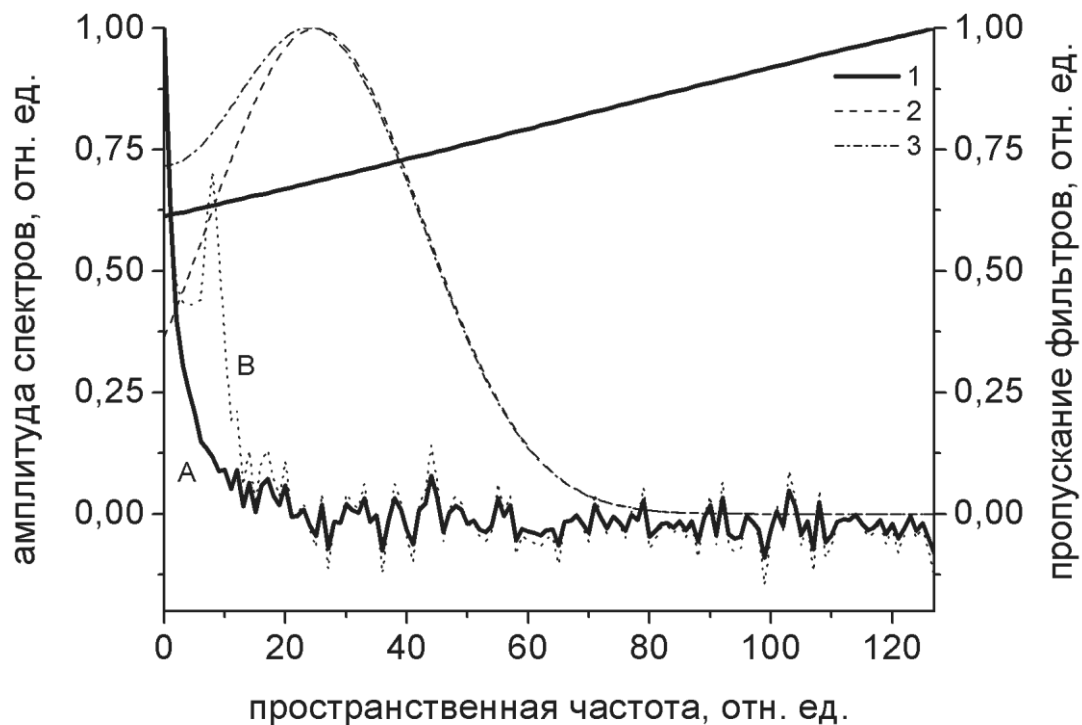


Рис. 3. Спектры реализаций (А – экспоненциальный, В – спектр с максимумом, сдвинутым в область более высоких частот) и фильтры (1 – сигмоидальный, 2 – гауссообразный, 3 – сумма двух гауссоид)

Поскольку нестационарность ограниченной реализации случайного процесса по матожиданию обусловлена преимущественно низкочастотными компонентами, моделировались три типа фильтров, описываемых приведёнными на рисунке 3 функциями, адекватно аппроксимирующими зависимость локальной дифракционной эффективности реальных высокочастотных голограмм Фурье для разных условий экспозиции и проявления [21]:

- сигмоидальной функцией (кривая 3);
- функцией Гаусса (кривая 4);
- суммой двух симметрично смещённых относительно нулевой частоты функций Гаусса (кривая 5).

Аналитические описания этих функций приведены в таблице 2.

Таблица 2

**Соответствие корреляционных функций параметрам,
определяющим сдвиг фильтрующей функции и изменение её формы**

Описание фильтра	Соответствие параметров фильтров и сечений корреляционных функций			
Сигмоидальный при сдвиге максимума $f(\nu) = \left(1 + \exp \frac{(-70+\nu)}{93}\right)^{-1} + 0.2 \cdot \left(1 - \frac{p}{3.5}\right)$	Рис. 3а кривая 1; фильтра нет	Рис. 3а кривая 2; p=0	Рис. 3а кривая 3; p=3	Рис. 3а кривая 4; p=6
Гауссообразный при сдвиге максимума $f(\nu) = \exp\left(-\frac{(\nu-25-1.1\nu_0)^2}{620}\right)$	Рис. 3а кривая 1; фильтра нет	Рис. 3а кривая 5; $\nu_0=0$	Рис. 3а кривая 6; $\nu_0=3$	Рис. 3а кривая 7; $\nu_0=6$
Сигмоидальный при изменении полуширины $f(\nu) = \left(1 + \exp \frac{(-70+\nu)}{93-6p}\right)^{-1} + 0.2$	Рис. 3б кривая 1; фильтра нет	Рис. 3б кривая 2; p=0	Рис. 3б кривая 3; p=3	Рис. 3б кривая 4; p=6
Гауссообразный при изменении полуширины $f(\nu) = \exp\left(-\frac{(\nu-25)^2}{620-42\nu_0}\right)$	Рис. 3б кривая 1; фильтра нет	Рис. 3б кривая 5; $\nu_0=0$	Рис. 3б кривая 6; $\nu_0=3$	Рис. 3б кривая 7; $\nu_0=6$

Причины выбора таких функций следующие:

1. Сигмоидальная функция часто используется в теории искусственных нейронных сетей в качестве активационной функции искусственного нейрона. Такая функция может быть реализована при выборе достаточно высокой частоты равенства амплитуд опорного и сигнального пучков при записи голограммы. На рисунке 2 параметры этого фильтра исходя из наглядности результатов выбраны так, чтобы при их изменении уменьшение радиуса корреляции происходило не слишком быстро.

2. Функции Гаусса и сумма двух функций Гаусса адекватно аппроксимируют передаточные характеристики реальных голограмм Фурье [21] и имеют практическое значение с точки зрения возможности реализации искусственных нейронных сетей с радиальными базисными функциями.

Для оценки влияния низких частот на стационарность по матожиданию моделировались методы настройки фильтров, адекватные реализуемым при записи голограмм Фурье:

- сдвиг частоты максимума пропускания фильтров в высокочастотную область при сохранении формы фильтрующей функции – этот метод моделирует изменение отношения амплитуд опорного и сигнального пучков при записи голограммы Фурье;

- изменение формы фильтрующей функции при постоянной частоте максимума – метод моделирует влияние параметров и условий обработки регистрирующей среды.

На рисунке 4 приведены сечения корреляционных функций для разных фильтров, описание которых, как и соответствие кривых на рисунке 4 параметрам, определяющим сдвиг фильтрующей функции и изменение её формы, приведены в таблице 2.

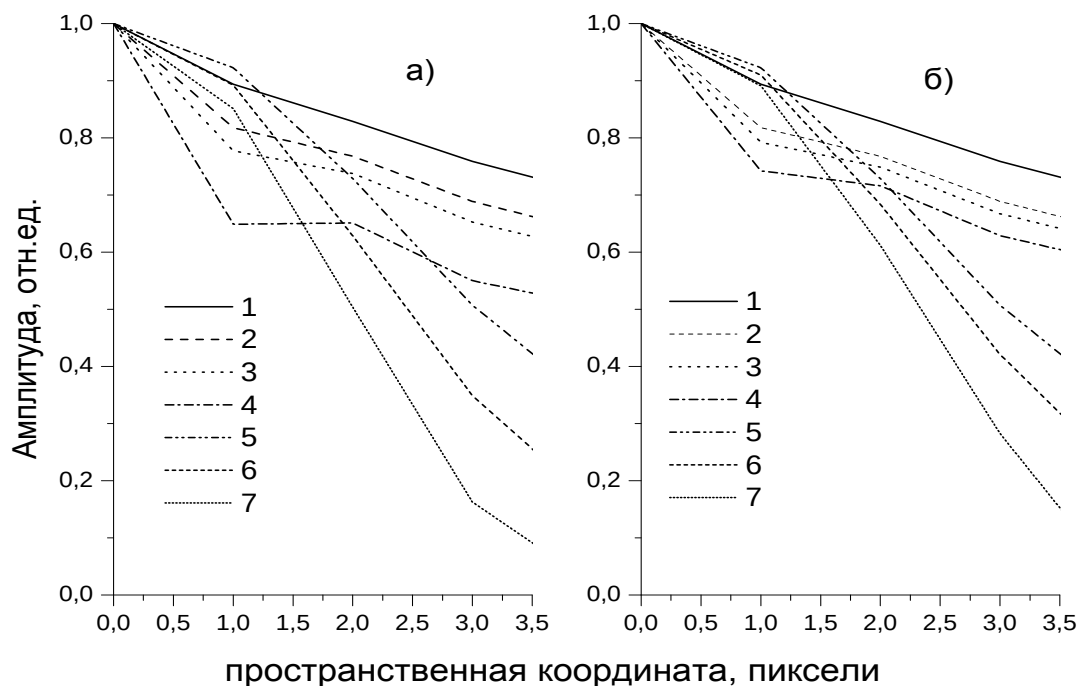


Рис. 4. Сечения корреляционных функций: а) при сдвиге максимума функции фильтра в область высоких частот; б) при сужении функции фильтра; 1 – до фильтрации, 2, 3, 4 – после фильтрации сигмоидальным фильтром; 5, 6, 7 – после фильтрации гауссообразным фильтром

При фильтрации сигмоидальным фильтром происходит сужение глобального максимума корреляционной функции. При фильтрациях, описываемых функцией Гаусса и суммой двух функций Гаусса, происходит изменение формы глобального максимума – увеличение радиуса (уплощение) вершины за счёт уменьшения пропускания в высокочастотной области при уменьшении радиуса у основания как результат уменьшения пропускания в низкочастотной области.

При фильтрации сигмоидальным фильтром радиус корреляции убывает быстрее, чем при фильтрации гауссообразным. Причина этого – в большем ослаблении высоких частот гауссообразным фильтром. При обоих типах фильтрации с увеличением величины параметров уменьшается радиус корреляционной функции, так как уменьшается пропускание в области низких частот (при сдвиге фильтрующей функции в высокочастотную область), а также в области высоких и низких частот (при сужении фильтрующей функции). Форма корреляционных функций при сужении и сдвиге максимума фильтра изменяется схожим образом.

На рисунке 5 приведены зависимости оценок стационарности от изменения функции фильтра (сужения фильтрующей функции – кривые 1–3) и от сдвига максимума функции фильтра (кривые 4–6) для двух видов спектров.

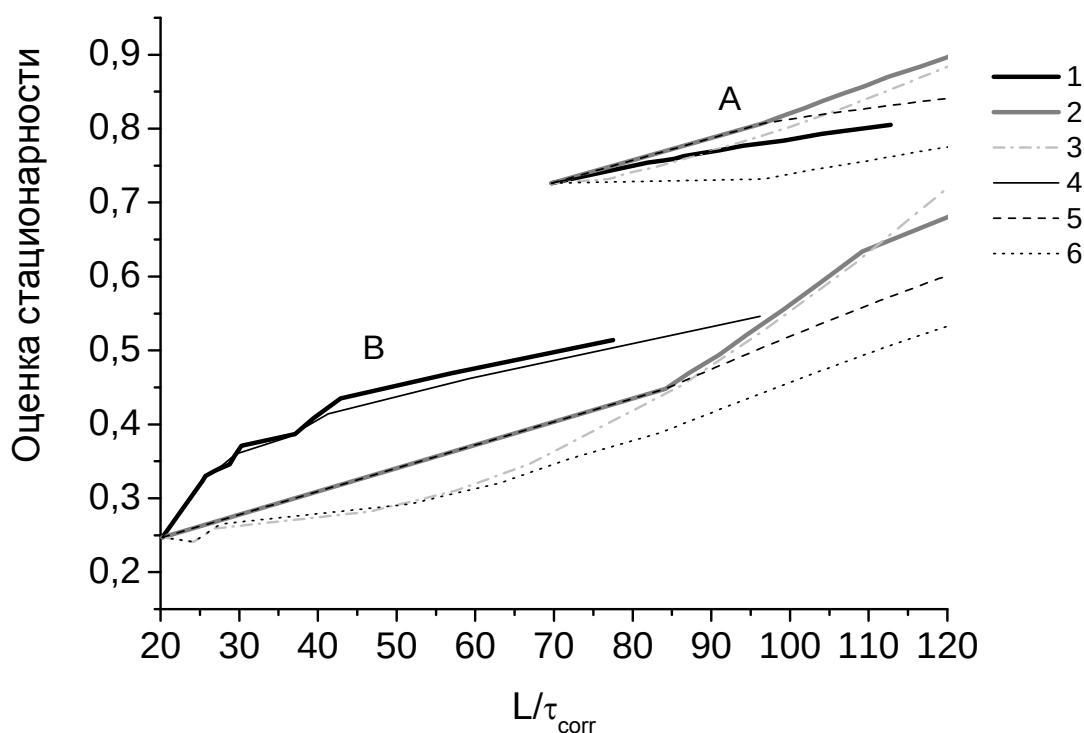


Рис. 5. Оценки стационарности для спектра с локальным максимумом (семейство кривых А) и для экспоненциального спектра (семейство кривых В) при следующих видах фильтрации:

- а) фильтрации путём сужения функции фильтра: 1 – при фильтрации сигмоидальным фильтром; 2 – при фильтрации гауссообразным фильтром; 3 – при фильтрации фильтром – суммы двух гауссоид;
- б) фильтрации путём смещения максимума функции фильтра: 4 – при фильтрации сигмоидальным фильтром; 5 – при фильтрации гауссообразным фильтром; 6 – при фильтрации фильтром – суммой двух гауссоид

Оценки стационарности вычислялись по формуле:

$$V_M\left(\frac{L}{\tau_{\text{corr}}}\right) = 1 - \left(\frac{(M_{\text{left}} - M_{\text{right}})^2}{D}\right),$$

где L – длина реализации; τ_{corr} – радиус корреляции; M_{left} и M_{right} – матожидание левой и правой частей реализации соответственно, полученные в результате деления реализации пополам; D – дисперсия реализации. В качестве аргумента функции оценки выбрано отношение длины реализации к радиусу корреляции по уровню 0,5. Это отношение, называемое иначе обобщённой пространственной частотой, является интегральной оценкой, связывающей характеристику процесса в пространственной области (длину реализации) со спектральной характеристикой через радиус корреляции. Так как взята длина реализации 256 пикселей, то рассмотрение производилось до величины отношения $L/\tau_{\text{corr}}=120$, поскольку в этих пределах ещё сохраняется внутренняя коррелированность процесса.

Приведённые на рисунке 5 зависимости позволяют оценить динамику изменения стационарности и при других значениях сечений корреляционной функции. Различия в начальных оценках реализаций с двумя типами спектров (до фильтрации) обусловлены разными радиусами корреляции. Оценки стационарности реализации в зависимости от сужения функции фильтра возрастают быстрее, чем в зависимости от сдвига максимума функции фильтра. Это обусловлено тем, что при изменении фильтрации путём сужения фильтрующей функции уменьшается удельный вес не только низких, но и высоких частот. Соответственно, радиус корреляции уменьшается медленнее, чем при изменении фильтрации путём сдвига функции фильтра в область высоких частот. Для спектра с локальным максимумом наибольшую оценку стационарности даёт фильтр с самым низким пропусканием в области низких частот – гауссообразный фильтр. Наименьшая оценка фильтрации – суммой двух функций Гаусса, так как этот фильтр имеет самое высокое пропускание в области низких частот. При фильтрации гауссообразным фильтром и фильтром-суммой двух функций Гаусса сужение фильтрующей функции более эффективно, чем сдвиг в высокочастотную область.

Наибольшее возрастание оценки стационарности для реализации с экспоненциальным спектром даёт фильтрация сигмоидальным фильтром. Это обусловлено тем, что доля низких частот в спектре с локальным максимумом уменьшается медленнее (большая часть веса низких частот приходится на локальный максимум), и, следовательно, радиус корреляции также уменьшается медленнее. Такой результат подтверждает предположение о преобладающем влиянии на стационарность по матожиданию именно низких частот.

Влияние режекции низких частот на оценку стационарности по матожиданию

Для проверки гипотезы о возрастании оценки стационарности при равенстве длины реализации целому числу периодов самой низкой частоты $T_{\min}$ моделировалась фильтрация вышеописанными фильтрами с дополнительной режекцией низких частот. Кривые оценок стационарности для такой фильтрации, приведённые на рисунке 6, имеют участок осцилляций, которые появились в результате последовательного удаления низких частот из спектра. Когда период самой низкой частоты кратен длине реализации (величина $L/T_{\min}$ целая), оценка растёт, если же $L/T_{\min}$ дробное, то, пока величина $L/T_{\min}$ мала, оценка снижается. По мере возрастания числа периодов в длине реализации эти осцилляции затухают.

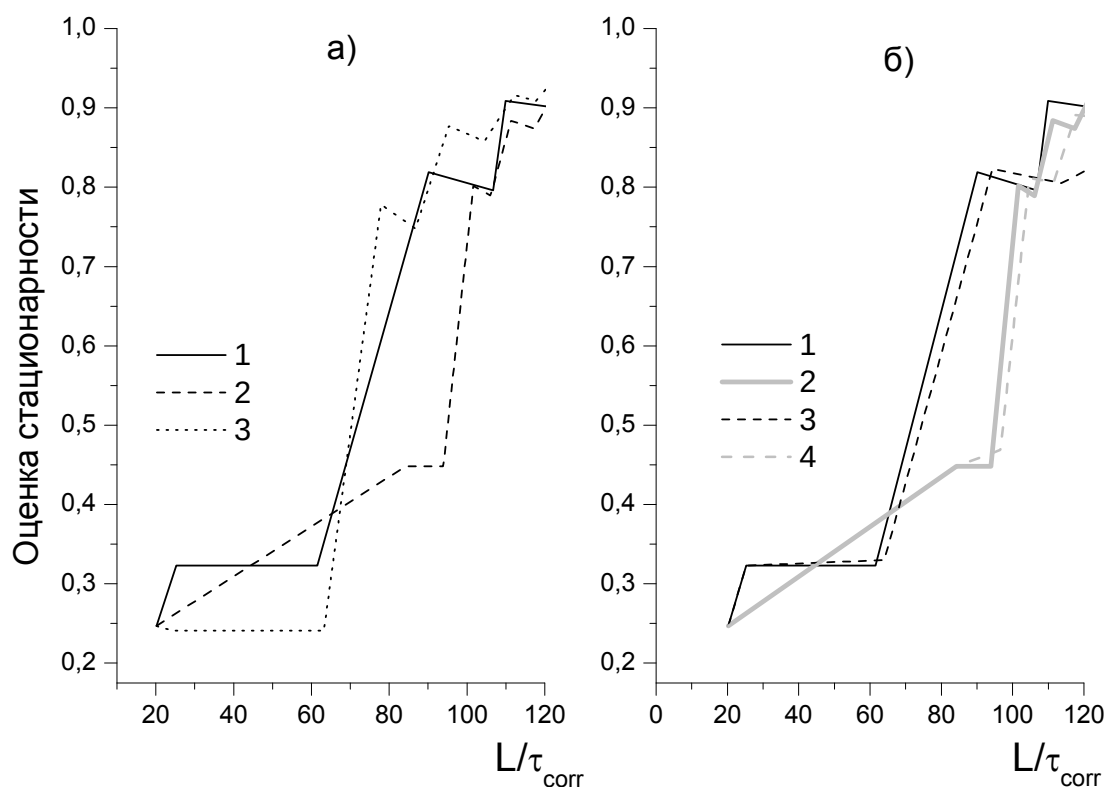


Рис. 6. Оценки стационарности для экспоненциального спектра при следующих видах фильтрации: а) при режекции: 1 – при фильтрации сигмоидальным фильтром; 2 – при фильтрации гауссообразным фильтром; 3 – при фильтрации фильтром-суммой двух функций Гаусса; б) при режекции и режекции с сужением фильтрующей функции: 1, 3 – режекция сигмоидального и гауссообразного фильтров соответственно; 2, 4 – режекция с изменением формы сигмоидального и гауссообразного фильтров соответственно

Оценки стационарности, приведённые на рисунке 6а (при режекции), лежат несколько выше, чем для фильтров 3–5 (рис. 3) при фильтрации без режекции. Действительно, при ослаблении низких частот фильтрами 3–5 (рис. 3) происходило лишь уменьшение удельного веса низких частот в спектре, но не полное их удаление. Несмотря на существенное ослабление, низкие частоты, в силу их большого удельного веса в спектре, всё равно вносили свой вклад в матожидание реализации. Здесь отношение длины реализации к периоду наименьшей частоты было постоянно, однако оценка стационарности по мере ослабления низких частот увеличивается. Поэтому кратность периода низкой частоты длине реализации в случае фильтра, только ослабляющего низкие частоты, не столь важна. Результаты, приведённые на рис. 6а, ещё раз подтверждают определяющую роль в оценке стационарности по матожиданию именно низких частот.

Поскольку возрастание оценок стационарности дали фильтрацию как путём режекции низких частот, так и путём сужения фильтрующей функции, то для достижения большей стационарности исследовано совмещение этих двух способов высокочастотной фильтрации – режекция низких частот и изменение формы фильтра. Оценки стационарности, полученные при фильтрации таким способом, также имеют осциллирующий характер, что является влиянием режекции низких частот.

На рисунке 6б оценка стационарности, полученная только при режекции низких частот, выше, чем оценка стационарности, полученная при комбинации режекции с сужением функции фильтра. Это обусловлено тем, что при фильтрации сигмоидальным фильтром разница в оценках стационарности происходит лишь из-за различия величин пропусканий в области низких частот (при фильтрации путём совмещения режекции и изменения формы функции фильтра пропускание уменьшается быстрее, следовательно, быстрее уменьшается радиус корреляции). При фильтрации гауссообразным фильтром и фильтром-суммой двух функций Гаусса разница оценок обусловлена тем, что при совмещении двух типов фильтрации имеют место два разнонаправленных процесса: уменьшение радиуса корреляции при уменьшении пропускания низких частот и увеличение радиуса корреляции при ослаблении высоких частот. Уменьшение веса низких частот в спектре оказывает большее влияние на оценку стационарности, радиус корреляции уменьшается быстрее, кривые за счёт этого как бы «растягиваются» вправо.

Заключение

Таким образом, ограниченность динамического диапазона голографических регистрирующих сред, используемых для реализации матрицы связей нейронных слоёв, имеет следствием повышение стационарности формируемых сетью гипотез по сравнению с характеристиками ПВР условий задачи. Низкочастотные голограммы влияют на повышение стационарности по дисперсии. Высокочастотные и полосовые голограммы ведут, главным образом, к повышению оценок стационарности по матожиданию. Эффективность повышения оценок стационарности по матожиданию определяется в первую очередь степенью ослабления фильтром низких частот.

Список литературы

1. Арбиб М. Метафорический мозг. – М.: Мир, 1976.
2. Прибрам К. Языки мозга. – М.: Прогресс, 1975.
3. Кузнецов, О.П. Неклассические парадигмы в ИИ // Известия АН. Сер. Теория и системы управления. – 1995. – № 5. – С. 3–23.
4. Голицын Г.А., Фоминых И.Б. Нейронные сети и экспертные системы: перспективы интеграции // Новости искусственного интеллекта. – 1996. – № 4.
5. Отражение образного мышления и интуиции специалиста в системах искусственного интеллекта // Новости искусственного интеллекта. – 1998. – № 1. – С. 22–136.
6. Дискуссионная трибуна. Отражение образного мышления и интуиции специалиста в системах искусственного интеллекта // Новости искусственного интеллекта. – 1998. – № 3. – С. 64–100.
7. О технологии решения творческих задач: сб. тр. / VIII Национальная конференция по искусственному интеллекту «КИИ-2002». – М.: Физматлит, 2002. – Т. 1.
8. Леутин В.П., Николаева Е.И. Функциональная асимметрия мозга. Мифы и реальность. – СПб.: Речь, 2005. – 368 с.
9. Динамика мозга в восприятии и сознании: творческая роль хаоса: сб. науч. тр. / Синергетика и психология. Вып. 3. Когнитивные процессы. – Когито-Центр, 2004. – С. 13–28.
10. Методы нелинейной динамики в когнитивной науке : сб. науч. тр. / Синергетика и психология. Вып. 3. Когнитивные процессы. – Когито-Центр, 2004. – С. 29–48.
11. Сознание: Хаотическое и странно-аттракторное : сб. науч. тр. / Синергетика и психология. Вып. 3. Когнитивные процессы. – Когито-Центр, 2004. – С. 49–60.
12. Павлов А.В. О возможности реализации абдуктивного и индуктивного выводов на нейросетях с матрицей связей, формируемой методом голографии Фурье // Искусственный интеллект и принятие решений. – 2010.
13. Интегрированные модели и мягкие вычисления в искусственном интеллекте: IV Междунар. науч.-практич. конф., 28–30 мая 2007 г., Коломна. – М.: Физматлит, 2007. – С. 282–290.

14. О возможности моделирования творческого мышления методом голографии Фурье : сб. науч. тр. / XI Национальная конф. по искусственному интеллекту. – М.: URSS. – Т. 1. – С. 384–392.
15. Информационные системы и технологии. Теория и практика: сб. науч. тр. / редкол.: А.Н. Берёза [и др.]. – Шахты: ГОУ ВПО «ЮРГУЭС», 2009. – С. 91–100.
16. Павлов А.В. Возможности ассоциативной обработки информации, реализуемые методом Фурье-голографии // Новости искусственного интеллекта. – 2006. – № 2. – С. 41–56.
17. Павлов А.В. О применимости модели линейной регрессии к описанию Фурье-голографии // Оптика и спектроскопия. – 2005. – Т. 98, № 6. – С. 1033–1037.
18. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и инженерные приложения. – М.: Высш. шк., 2000.
19. Вентцель А.Д. Курс теории случайных процессов. – М.: Наука, 1975.
20. Grimmet G.R. Probability and Random Processes / G.R. Grimmet, D.R. Sterzaker. – Oxford Sc. Publ., Clarendon Press, 1992.
21. Алексеев А.М., Павлов А.В. Об аппроксимации передаточной функции 4f схемы голографии Фурье при реализации нечётко-значимых логик // Оптика и спектроскопия. – 2009. – Т. 107. – № 6. – С. 1038–1043.
22. Павлов А.В. Реализация модели линейного предсказателя методом Фурье-голографии // Оптический журнал. – 2005. – Т. 72, № 2. – С. 43–47.

УДК 535.4:159.9

В.В. Орлов

**АССОЦИАТИВНЫЕ СВЯЗИ МЕЖДУ СЛОВАМИ,
ХРАНЯЩИМИСЯ В ПАМЯТИ ЧЕЛОВЕКА,
СОГЛАСНО ГОЛОГРАФИЧЕСКОЙ МОДЕЛИ ПАМЯТИ***

* Работа выполнена при поддержке РФФИ, грант 09-01-00165-а.

Рассмотрена голографическая модель памяти человека, согласно которой вербальная и образная информации хранятся в памяти в виде объёмных наложенных голограмм, у которых образы играют роль объектных волн, вербальная информация – роль опорных волн. Ранее было установлено, что нелинейность процесса записи голограмм приводит к возникновению ассоциативной связи между вербальными элементами информации, т.е. между словами и мыслями человека. В настоящей работе показано, что данная ассоциативная связь возникает также в результате двукратной дифракции волн в объёме голограмм. Для возникновения этого вида ассоциативной связи необходимо и достаточно нарушение взаимной ортогональности объектных волн голограмм. Найдено отношение сил двух данных видов ассоциативной связи. Показано, что их силы одинаково зависят от степени взаимной ортогональности объектных волн голограмм.

Ключевые слова: голографическая модель памяти, ассоциативная память, наложенные голограммы, гельштат-психология, вербальная информация.

Введение. Современные компьютеры превосходят по быстродействию мозг человека в миллионы раз, память компьютера, способная хранить и почти с абсолютной точностью воспроизводить огромные массивы информации, также намного превосходит возможности памяти человека. Тем не менее, компьютер не способен не только мыслить подобно человеку, но и воспроизводить функции мозга животного. Широко распространено мнение, что уникальные способности мозга обусловлены тем, что информация в нём, в отличие от компьютера, обрабатывается параллельно, а не последовательно. Параллельные информационные процессы имеют место также в голографии, например, когда с голограммы восстанавливается изображение. Как мозг, так и голограмма имеют свойства ассоциативности, когда по фрагменту образа восстанавливается весь образ, и дистрибутивности, когда информация о каждом образе распределена по всему объёму носителя информации. Общность свойств мозга и голограммы привела к гипотезе о том, что в основе памяти человека лежит голографический принцип записи и восстановления волнового поля [1]. В соответствии с этой гипотезой были предложены голографические модели памяти человека [2]. Нами было отмечено, что если память человека действительно функционирует подобно голограмме, то в основе памяти должен лежать принцип записи и восстановления наложенных, или, как их ещё называют, мультиплексных, голограмм, т.е. голограмм, записанных на одном участке регистрирующей среды. Была предложена голографическая модель памяти

[3], основанная на свойствах объёмных наложенных голограмм. Согласно этой модели, вербальная и образная информации хранятся в памяти в виде наложенных голограмм, записанных методом углового мультиплексирования, у которых образы играют роль объектных волн, вербальная информация – роль опорных волн. Было показано, что свойства такой модели памяти соответствуют положениям гельштат-психологии.

Ассоциативная связь между вербальными элементами информации. Ранее был рассмотрен механизм возникновения ассоциативной связи между вербальными элементами информации, хранящимися в памяти человека, т.е. между словами и мыслями человека, возникающий при условии квадратичной нелинейности процесса записи наложенных голограмм [4]. В настоящей работе рассмотрен второй механизм возникновения ассоциативной связи между вербальными элементами информации. Этот механизм не требует нелинейности при записи голограмм, достаточно нарушения взаимной ортогональности объектных волн голограмм. Показано, что в этом случае ассоциативная связь возникает в результате двукратной дифракции волн в объёме голограмм.

Пусть из памяти человека восстанавливается вербальный элемент информации, т.е. слово, которому соответствует опорная волна R_n . Распространяясь в объёме наложенных голограмм, опорная волна в результате первого акта дифракции восстанавливает соответствующую ей объектную волну O_n , представляющую собой образ, названием которого служит данное слово. Волна O_n в следующем, втором, акте дифракции восстанавливает

опорные волны других наложенных голограмм $\sum_{m=1} R_m$, т.е. другие вербальные элементы, при условии, если она не ортогональна по отношению к объектным волнам этих голограмм. Последовательность двух данных актов дифракции описывается следующим выражением:

$$R_n \rightarrow O_n \rightarrow \sum_{m=1} R_m,$$

где символ $\rightarrow$ означает один акт дифракции.

Для нахождения амплитуд восстановленных опорных волн воспользуемся выражением (12) из работы [5]:

$$\vec{R}_n(z) = \sum_{m=1}^K \sum_{l=1}^M S_{mn}^* \cos(Pz\sqrt{\tau_m}) S_{ml} \vec{R}_l + \sum_{m=K+1}^M \sum_{l=1}^M S_{mn}^* S_{ml} \vec{R}_l, \quad (1)$$

которое описывает зависимость комплексных амплитуд $\vec{R}_n(z)$, восстановленных опорных волн наложенных голограмм от толщины голограмм z при падении на голограммы опорной волны n -й наложенной голограммы. Положив в (1)

$$\cos x \approx 1 - \frac{x^2}{2},$$

что соответствует учёту первых двух актов дифракции волн в объёме голограмм, получим:

$$\vec{R}_{rn}(z) = \vec{R}_n - \frac{1}{2} P^2 z^2 Q \sum_{l=1}^M G_{ln} \vec{R}_l, \quad (2)$$

где P – параметр, зависящий от условий записи голограмм, Q – интенсивность опорных волн; G_{ln} – скалярное произведение l -й и n -й объектных волн; M – число наложенных голограмм; $\vec{R}_l$ – вектор комплексных амплитуд l -й опорной волны. Из (2) следует, что при освещении наложенных голограмм n -й опорной волной восстанавливаются опорные волны других наложенных голограмм, амплитуды которых пропорциональны скалярному произведению объектных волн этих голограмм и объектной волны n -й голограммы. Для голографической модели памяти это означает, что возникает ассоциативная связь между вербальными элементами информации, хранящимися в памяти, сила которой пропорциональна скалярному произведению образов, соответствующих вербальным элементам информации.

Итак, для рассматриваемой голографической модели памяти существует два механизма формирования ассоциативной связи между вербальными элементами информации, хранящимися в памяти. Один механизм обусловлен нелинейной записью голограмм [3], второй – двукратной дифракцией волн в объёме голограмм. Оба механизма формирования ассоциативной связи работают «автоматически», без участия какой-либо «управляющей программы». Ассоциативная связь возникает между элементами вербальной информации, запоминаемыми в разные моменты времени. Сравним силы этих ассоциативных связей, приняв в качестве критерия силы связи дифракционную эффективность, с которой восстанавливается опорная волна, соответствующая вербальному элементу, восстановленному из памяти по ассоциации, с исходным вербальным элементом. В данном случае дифракционная эффективность определяется как отношение интенсивности восстановленной опорной волны к интенсивности опорной волны исходного вербального элемента.

Для ассоциативной связи, возникающей в результате квадратичной нелинейности, получим следующее значение дифракционной эффективности:

$$\eta_{2N}(z) = \frac{1}{\cos^2 \theta} \kappa^2 \varepsilon_0 \chi_2^2 t^4 Q^2 |G_{nl}|^2 z^2. \quad (3)$$

Для ассоциативной связи, возникающей в результате двух актов дифракции, дифракционная эффективность описывается следующим выражением:

$$\eta_{2D}(z) = \frac{1}{64 \cos^4 \theta} \kappa^4 \varepsilon_0^2 \chi_c^4 t^4 Q^2 |G_{nl}|^2 z^4, \quad (4)$$

где θ – углы падения опорных и объектных волн (схема записи голограмм симметричная), κ – волновое число. Из (9), (10) следует, что в обоих случаях сила ассоциативной связи между вербальными элементами информации пропорциональна квадрату модуля скалярного произведения G_{nl} образов, соответствующих данным вербальным элементам.

При формировании ассоциативного отклика голограмма, освещаемая опорной волной, одновременно восстанавливает свою объектную волну. Поэтому целесообразно сравнить дифракционные эффективности (3), (4) с дифракционной эффективностью $\eta(z)$, с которой наложенная голограмма восстанавливает объектную волну. Для этого воспользуемся выражением (24) из работы [5] и выразим дифракционные эффективности (3), (4) через дифракционную эффективность $\eta(z)$ наложенной голограммы:

$$\eta_{2N}(z) = 4 \left(\frac{\chi_2}{\chi_c} \right)^2 t^2 Q G_{nn} \frac{|G_{nl}|^2}{G_{nn}^2} \eta(z), \quad (5)$$

$$\eta_{2D}(z) = \frac{1}{4} \frac{|G_{nl}|}{G_{nn}^2} \eta^2(z). \quad (6)$$

Данные выражения позволяют определить отношения интенсивностей восстановленных голографической моделью памяти волны образа, названием которого служит исходный вербальный элемент информации, и волны вербального элемента информации, восстановленного по ассоциации с исходным вербальным элементом.

При записи большого числа наложенных голограмм их дифракционная эффективность $\eta(z)$ мала, её величина принимает значения, меньшие 0,001 [6]. В этом случае, согласно (6), интенсивность ассоциативного отклика, при условии $\frac{|G_{nl}|}{G_{nn}^2} \approx 1$, более чем в 1000 раз меньше интенсивности

образа. Волна образа может состоять из многих тысяч компонент, в то время как волна вербального элемента информации состоит из одной компоненты. Поэтому интенсивность одной компоненты волны вербального элемента информации может быть больше интенсивности любой компоненты волны образа, что способствует фиксации внимания человека на вербальном элементе информации.

Сравним интенсивность ассоциативного отклика, обусловленного двукратной дифракцией волн, с интенсивностью ассоциативного отклика, обусловленного квадратичной нелинейностью регистрирующей среды. Для этого найдём отношение их дифракционных интенсивностей (6) и (5):

$$\frac{\eta_{2D}(z)}{\eta_{2N}(z)} = \frac{\pi^2}{16 \cos^2 \theta} \left(\frac{\chi_c^2}{\chi_2} \right)^2 \varepsilon_0 \left(\frac{z}{\lambda} \right)^2, \quad (7)$$

где λ – длина волны света.

Данное равенство определяет зависимость отношения интенсивностей двух видов ассоциативной связи от параметров регистрирующей среды χ_c и χ_2 , толщины голограммы z и длины волны λ используемого излучения. Из (7) следует, что увеличение толщины голограммы приводит к увеличению интенсивности ассоциативной связи, обусловленной двукратной дифракцией волн по сравнению с интенсивностью ассоциативного отклика, обусловленного квадратичной нелинейностью регистрирующей среды.

Чтобы получить объективное доказательство существования данной ассоциативной связи слов, студентам было предложено сгруппировать слова из предложенного им списка слов в пары слов. Как пример, следуя которому должны быть образованы пары слов, им была представлена пара слов «часы – время». Был предложен следующий список слов:

Бык	Мама
Ветка	Море
Дерево	Разность
Дрова	Ребёнок
Замок	Рога
Квартира	Сумма
Ключ	Тёплый
Корабль	Холодный
Костёр	Шкаф

В эксперименте участвовало 24 студента, выполняющих задание в индивидуальном порядке. Результаты эксперимента представлены в следующей таблице, где на пересечении строк и столбцов показано число студентов, образовавших данную пару слов, нормированное на число студентов, участвовавших в эксперименте.

	Рога	Дерево	Костёр	Ключ	Шкаф	Море	Ребёнок	Сумма	Холодный
Бык	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Ветка	0,00	0,92	0,08	0,00	0,00	0,00	0,00	0,00	0,00
Дрова	0,00	0,08	0,88	0,04	0,00	0,00	0,00	0,00	0,00
Замок	0,00	0,00	0,00	0,79	0,21	0,00	0,00	0,00	0,00
Квартира	0,00	0,00	0,00	0,21	0,79	0,00	0,00	0,00	0,00
Корабль	0,00	0,00	0,00	0,00	0,00	0,96	0,00	0,00	0,04
Мама	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00	0,00
Разность	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	0,00
Тёплый	0,00	0,00	0,04	0,00	0,00	0,04	0,00	0,00	0,92

Как видно из таблицы, независимо от индивидуальных особенностей человека в его памяти существует ассоциативная связь между определёнными словами.

Предложенная голографическая модель памяти объясняет возникновение такой ассоциативной связи слов. Данная ассоциативная связь играет важную роль в мыслительной деятельности человека, позволяя создать цепочку слов, ведущую к решению поставленной задачи. Например, предположим, что человек оказался в холодное время года в лесу. Чтобы согреться, он может разжечь костёр, для чего следует собрать упавшие ветки или сломать ветки деревьев. Цепочка слов, ассоциированных друг с другом и порождающих соответствующие им мысли, может быть такой: холодный → тёплый → костёр → дрова → деревья → ветки.

Модель памяти человека, в основу которой положен принцип записи и восстановления объёмных наложенных голограмм, объясняет возникновение ассоциативной связи между словами и мыслями, хранящимися в памяти человека. Данная модель, как было показано ранее, соответствует основным положениям гештальт-психологии. Вопрос, насколько предложенная модель памяти соответствует реальности, остаётся открытым до получения новых данных о работе мозга человека.

Список литературы

1. Pribram K.H. The newophysioleoy of remembering // Scientific American. – 1969. – V. 220. – P. 73–86.
2. Кузнецов О.П., Марковский Л.Б., Шипилина Л.Б. Голографические механизмы обработки образной информации. – М.: Институт проблем управления им. В.А. Трапезникова РАН, 2007.
3. Орлов В.В. Подобие свойств объёмных наложенных голограмм свойствам памяти человека // Оптический журнал. – 2006. – № 9. – С. 83–86.
4. Орлов В.В. Формирование ассоциативных связей в вербальной части памяти человека согласно голографической модели памяти // Труды Конгресса по интеллектуальным системам и информационным технологиям «AIS-IT'09». – М.: Физматлит, 2009. – Т. 2. – 568 с. – С. 200–202.
5. Орлов В.В. Модовая теория объёмных наложенных голограмм. Волновое поле восстановленной голограммы // Оптика и спектроскопия. – 2002. – Т. 92, № 5. – С. 862–866.
6. Coufal H.J. Psaltis D., Sincerbos G.T., Holographic Data Storage. – Berlin: Springer-Verlage, 2000. – 490 с.

РАЗДЕЛ 4

СИСТЕМЫ АВТОМАТИЗАЦИИ ПРОЕКТИРОВАНИЯ

УДК 004.4'423, 004.4'424

Л.А. Гладков^{*}, А.Е. Лисовцова

ПРИМЕНЕНИЕ НЕЧЁТКИХ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ С ВЕЩЕСТВЕННЫМ КОДИРОВАНИЕМ ДЛЯ РЕШЕНИЯ ЗАДАЧ ОПТИМИЗАЦИИ^{**}

Работа эволюционных алгоритмов основывается на принципах, которые заимствованы из биологии и генетики. Основная идея состоит в создании популяции решений (хромосом). Любая хромосома есть возможное решение рассматриваемой оптимизационной задачи. Значение целевой функции или функции приспособленности решения показывает, насколько хорошо оно подходит для решения задачи с учётом имеющихся ограничений. Хромосома состоит из конечного числа генов, представляя генотип объекта, т.е. совокупность его наследственных признаков. К популяции применяются различные биологические операторы: кроссинговера, мутации, инверсии и др. Популяция постоянно обновляется при помощи генерации новых особей и уничтожения старых (менее приспособленных) индивидов.

Основное преимущество генетических алгоритмов заключается в том, что они позволяют решать комбинаторные задачи большой размерности и преодолевать возможные барьеры локальных оптимумов. Генетические алгоритмы позволяют также работать с функциями, имеющими сложный ландшафт поверхности, а также на поверхностях с разрывами, плоскими участками и другими сложными неупорядоченными формами.

В зависимости от применяемого способа кодирования различают генетические алгоритмы с двоичным, десятичным и т.д. кодированием. Также перспективным считается использование генетических алгоритмов с вещественным кодированием, где гены представлены вещественными числами.

^{*}347928, Таганрог, пер. Некрасовский, 44, ТТИ ЮФУ, leo@tsure.ru

^{**}Работа выполнена при поддержке РФФИ, грант № 08-01-00473, и программы развития научного потенциала высшей школы 2009–2010 гг. (РНП. 2.1.2.1652).

Вещественные генетические алгоритмы нашли широкое применение при работе с оптимизационными задачами в непрерывных пространствах. В этом случае хромосома представлена вектором вещественных чисел, моделирующих свойства объекта оптимизации. То есть каждый ген отвечает за конкретное свойство объекта. Точность вычислений в этом случае зависит от разрядной сетки ЭВМ, на которой будут выполняться испытания. Можно выделить следующие преимущества, которые дают такие алгоритмы:

- непрерывные гены дают возможность осуществлять поиск в больших пространствах. При использовании двоичных генов уменьшается точность вычислений, что приводит к необходимости увеличивать размер популяции. Это в свою очередь скажется на увеличении времени и сокращении вычислений;
- использование вещественных генетических алгоритмов для представления решений удобно, поскольку близко к постановке большинства прикладных задач. Кроме того, отсутствие операций кодирования/декодирования, которые необходимы в двоичных генетических алгоритмах, повышает скорость работы алгоритма;
- генетические алгоритмы с вещественным кодированием просто использовать в задачах, где целевая функция и функции ограничений носят нелинейный характер. При этом основная трудность заключается в формировании начальной популяции, удовлетворяющей ограничениям задачи. Эксперименты показывают, что при формировании начальной популяции случайным образом в заданном диапазоне только около 10 % хромосом находятся в пределах области допустимых решений.

При решении задач оптимизации часто встречаются случаи, когда область допустимых решений, определяемая набором ограничений, разделена на несколько подобластей сложной формы, а число ограничений задачи велико. В этой ситуации использование традиционных алгоритмов оптимизации затруднительно.

Для кодирования признака, принимающего действительные значения в определённом диапазоне, используется специальный приём. Интервал допустимых значений признака x_i разбивают на участки с g_i требуемой точностью. Для преобразования целочисленного значения гена из множества $\{0, \dots, 2^N\}$ в вещественное число r_i из интервала пользуются формулой:

$$r_i = g_i \frac{(b_i - a_i)}{2^N - 1} + a_i,$$

где N – количество разрядов для кодирования битовой строки. Чаще всего используются значения $N = 8; 16; 32$.

При увеличении N пространство поиска катастрофически увеличивается. Рассмотрим такой пример. Пусть требуется найти минимум для 100 переменных, изменяющихся в интервале $[-500; 500]$, с точностью до шестого знака после запятой. При использовании генетических алгоритмов (ГА) с двоичным кодированием длина строки составит 3000 элементов, а пространство поиска – около 10 в степени 1000 . Применение вещественного кодирования повышает точность найденных решений и скорость нахождения оптимума.

Для решения таких задач в непрерывных пространствах был предложен генетический алгоритм с вещественным кодированием (англ.: Real-coded Genetic Algorithm, RGA) [1–3]. Основная идея RGA заключается в том, чтобы напрямую представлять гены в виде вещественных чисел, при этом генотип объекта становится идентичным его фенотипу. Хромосома представляет собой вектор вещественных чисел, точность решения при этом определяется не количеством разрядов для кодирования битовой строки, а ограничена возможностями ЭВМ, на которой реализуется вещественный ГА.

Для RGA стандартные операторы кроссинговера и мутации не подходят, т.к. алгоритм работает только с вещественными числами. По этой причине были разработаны специальные генетические операторы, которые описаны в работе [1]. Рассмотрим основные из них.

Пусть $C_1 = (C_1^1, C_2^1, \dots, C_n^1)$ и $C_2 = (C_1^2, C_2^2, \dots, C_n^2)$ – две хромосомы, выбранные оператором селекции для проведения кроссинговера.

Плоский кроссинговер. Создаётся потомок $H = (h_1, \dots, h_i, \dots, h_n)$, где $h_i, i = \overline{1, n}$ – случайное число на интервале $[C_i^1, C_i^2]$.

Арифметический кроссинговер. Создаются два потомка $H_1 = (h_1^1, \dots, h_n^1)$ и $H_2 = (h_1^2, \dots, h_n^2)$, причём $h_i^1 = \eta C_i^1 + (1 - \eta) C_i^2$, $h_i^2 = \eta C_i^2 + (1 - \eta) C_i^1$, где $i = \overline{1, n}$, $\eta \in [0, 1]$ – константа.

BLX α кроссинговер. Генерируется один потомок $H = (h_1, \dots, h_i, \dots, h_n)$, где h_i случайное число на интервале $[C_{\min} - \Delta\alpha, C_{\max} + \Delta\alpha]$, где $C_{\max} = \max(C_i^1, C_i^2)$, $C_{\min} = \min(C_i^1, C_i^2)$, $\Delta = C_{\max} - C_{\min}$, $i = \overline{1, n}$.

Линейный кроссинговер. Создаются три потомка, рассчитываемые по формулам:

$$h_i^1 = \frac{C_i^1 + C_i^2}{2}, h_i^2 = \frac{3C_i^1 + C_i^2}{2}, h_i^3 = \frac{-C_i^1 + 3C_i^2}{2}.$$

При селекции в линейном кроссинговере выбираются две особи с наибольшими приспособленностями.

В качестве оператора мутации наибольшее распространение получили случайная и неравномерная мутации Михалевича. При случайной мутации ген, подлежащий изменению, принимает случайное значение из интервала своего изменения. В неравномерной мутации значение гена после оператора мутации рассчитывается по формуле:

$$\tilde{C}_i = \begin{cases} C_i + \delta(t, b_i - C_i) & \text{при } \chi = 0 \\ C_i - \delta(t, C_i - a_i) & \text{при } \chi = 1 \end{cases},$$

$$\delta(t, y) = y \left(1 - r \left(1 - \frac{t}{\delta_{\max}} \right)^b \right),$$

где χ – целое случайное число, принимающее значение 0 или 1; $r \in [0, 1]$ – случайное вещественное число; $\delta_{\max}$ – максимальное количество эпох алгоритма; b_i – параметр, задаваемый исследователем.

В работе [1] приведены результаты экспериментов на тестовых функциях с использованием разных типов операторов скрещивания и мутации для ГА с вещественным кодированием (RGA). Например, классической тестовой функцией является N -мерная функция Розенброка, имеющая вид:

$$f(x_1, \dots, x_{N+1}) = \sum_{i=1}^N \left(100 \cdot (x_{i+1} - x_i^2)^2 + (1 - x_i)^2 \right) \rightarrow \min.$$

Оптимальные значения переменных при $f(x) = 0$ равны $x_i^{\oplus} = 1, i = \overline{1, N+1}$, для $x_i \in [-5, 5], i = \overline{1, N+1}$.

Кроме того, полученные результаты сравнивались с результатами работы ГА с двоичным (бинарным) кодированием. На основе данного сравнения сделан вывод о том, что в большинстве случаев генетический алгоритм с вещественным кодированием справляется с задачей нахождения оптимума лучше и быстрее, чем с двоичным кодированием. Самым эффективным оператором скрещивания признан BLX- α кроссинговер с $\alpha = 0,5$. Особенность данного оператора в том, что при скрещивании генов $C_i^1, C_i^2 (C_i^1 < C_i^2)$ значения потомка могут лежать в некоторой области, выходящей за границы значений этих генов на величину $\Delta\alpha$, т.е. $h_i \in [C_i^1 - \Delta\alpha, C_i^2 + \Delta\alpha]$. В других кроссинговерах, например в плоском или арифметическом, $h_i \in [C_i^1, C_i^2]$.

Авторами предлагается новый оператор кроссинговера, который является модификацией родительски-центрированного оператора кроссинговера (BLX или PBX), предложенного Шелманом и Сшифером. Его алгоритм можно описать следующим образом. Пусть $X = (x_1, \dots, x_n)$ и $Y = (y_1, \dots, y_n)$ ($x_i, y_i \in [a_i, b_i] \subset \mathbb{R}$, $i = 1, \dots, n$) – две хромосомы с вещественным кодированием, выбранные на основе стратегии выполнения оператора кроссинговера. В результате получаем потомков: $Z_1 = (z_1^1 \dots z_n^1)$ или $Z_2 = (z_1^2 \dots z_n^2)$, где z_i^1 – произвольным образом выбранное число на интервале $[l_i^1, u_i^1]$, причём

$$l_i^1 = \max\{a_i, x_i - I \cdot \alpha\} \text{ и } u_i^1 = \min\{b_i, x_i + I \cdot \alpha\},$$

а z_i^2 выбирается на интервале $[l_i^2, u_i^2]$ и

$$l_i^2 = \max\{a_i, y_i - I \cdot \alpha\} \text{ и } u_i^2 = \min\{b_i, y_i + I \cdot \alpha\}, \text{ где } I = |x_i - y_i|.$$

Можно выделить следующие основные свойства такого оператора:

- позволяет с большой вероятностью создавать потомство, расположенное в области допустимых решений задачи, близко к родителям. Это даёт эффективный способ для оптимизации вещественного параметра;
- степень разнообразия популяции решений после применения оператора можно легко изменять;
- оператор позволяет количественно оценивать разнообразие популяции и степень «близости» вновь полученных решений к родительским. Таким образом, использование данного оператора в RCGAs даёт возможность динамически, в процессе поиска, оценивать возможности адаптации решений-потомков.

Оператор кроссинговера на основе глобального поиска (ХНС) многократно применяется к выбранной паре родительских решений до тех пор, пока не будет получено заданное число n_{off} потомков. После этого лучший из потомков заменяет худшего родителя, если это не приводит к ухудшению целевой функции. Процесс повторяется в течении n_{it} итераций, после чего в популяцию возвращают двух исключённых текущих родителей. Очевидно, что при выполнении такого оператора должны быть заранее заданы значения n_{off} и n_{it} и выбрана начальная пара родителей (p_1, p_2) .

Предложенный оператор может быть использован при выполнении RCGA, т.к. для его реализации достаточно популяции минимального размера (не менее двух хромосом). При этом за счёт изменения правил отбора можно модифицировать данного оператора и преобразовать его в мультихромосомный. По результатам сравнения можно изменять текущую пару родителей в процессе выполнения оператора. При использовании решений с вещественным кодированием оценка качества решений производится на основе вероятности распределения популяции решений. Это позволяет регулировать процессы схождения и разнообразия решений в популяции.

Также применяется ряд методик и технологий организации процесса поиска, сочетающих достоинства локальных методов поиска и методов генетического поиска. К таким методикам относятся процедуры минимального разрыва поколений (Minimal generation gap -MGG) и обобщения поколений (G3-Generalizedgenerationgap).

Кратко идею процедуры обобщения поколений можно описать следующим образом:

- из популяции $P(t)$ выбирается лучший родитель, второй ($\mu-1$) родитель выбирается произвольно;
- создание потомства с использованием мультиродительского оператора кроссинговера;
- произвольный выбор двух элементов из популяции $P(t)$;
- создание новой подпопуляции из двух выбранных элементов и полученных потомков, выбор двух лучших решений и замена элементов.

В результате данной процедуры получаем новую подпопуляцию с улучшенным потомством, из которой впоследствии и будет выделен лучший вариант решения задачи.

В настоящее время разрабатывается ряд операторов кроссинговера, моделирующих адаптивное поведение, которое можем наблюдать при применении эволюционных стратегий и эволюционного программирования. Определяющим фактором для эволюции при этом является оценка удалённости между родителями в пространстве поиска.

Список литературы

1. Herrera F., Lozano M., Verdegay J.L. Tackling real-coded genetic algorithms: operators and tools for the behaviour analysis // Artificial Intelligence Review. – Vol. 12, No. 4. – 1998. – P. 265–319.
2. Michalewicz Z. Genetic Algorithms, Numerical Optimization and Constraints // Proceedings of the 6th International Conference on Genetic Algorithms, Pittsburgh, July 15–19, 1995. – P. 151–158.
3. Wright A. Genetic algorithms for real parameter optimization // Foundations of Genetic Algorithms. – 1991. – Vol. 1. – P. 205–218.

УДК 004.4'423

А.С. Стороженко, А.Г. Подстрелов

СИСТЕМА ОПТИМИЗАЦИИ МАТЕМАТИЧЕСКИХ ФУНКЦИЙ ДЛЯ ВЫЧИСЛЕНИЯ НА ЭВМ

В статье рассматриваются вопросы эффективной организации вычислений математических функций на современных ЭВМ, предлагается подход, обеспечивающий значительное сокращение времени вычисления, что подтверждается приведёнными результатами экспериментальных исследований.

Ключевые слова: математические функции, оптимизация вычислений, ЭВМ, трансляторы, ассемблер.

Введение. В настоящее время решение научных и инженерных задач сложно себе представить без применения ЭВМ. Это, как правило, задачи расчёта, моделирования, оптимизации, синтеза и ряд других проектных решений, для которых затруднительно применение ручных методов расчёта из-за, например, большой размерности задачи. Для разработки моделей и алгоритмов решения таких задач необходимы эффективные механизмы вычисления стандартных математических функций, таких, например, как линейные, циклические, тригонометрические и т.д. При конструировании моделей и алгоритмов, как правило, используется итерационный подход, заключающийся в выполнении следующих четырёх основных этапов: разработка и кодирование для реализации на ЭВМ математической модели, выбор и кодирование алгоритма решения, тестирование алгоритма, анализ полученных результатов. Итерационность подхода заключается в многократном прохождении этапов проектирования до получения необходимых результатов, что ведёт к значительному замедлению процесса конструирования.

Основные затраты времени на проектирование решаемой задачи приходятся на этап проведения экспериментальных исследований. Это связано со значительным временем, затрачиваемым на вычисление сконструированных математических моделей. Уменьшения времени проектирования можно добиться, сократив время вычисления математических моделей и как частное математических функций.

На сегодняшний день существует множество способов реализации вычисления математических функций на ЭВМ. Основными из них являются:

- использование специализированных программных сред (MatLab, MathCad и т.п.). Плюсом такого подхода является простота написания функций и вывода результата на экран; главным их минусом является медленное вычисление математических функций [8];

– использование для описания математических функций языков высокого уровня (C++, Delphi и т.п.). Плюсом данного подхода является ускорение вычисления математических функций за счёт частичной оптимизации представления математической функции для вычисления на ЭВМ; минусом – то, что пользователь должен обладать знаниями, умениями и опытом в области программирования и оптимизации кода. Также следует отметить, что даже хорошо оптимизированный код не позволит максимально быстро рассчитать математическую функцию;

– использование для описания математических функций языка ассемблер. Плюсом этого подхода является возможность организации максимально быстрого расчёта математических функций с полной оптимизацией его представления для ЭВМ; к минусу этого подхода можно отнести необходимость знания архитектурных и программных особенностей проектирования вычислений на конкретной ЭВМ.

Учитывая вышеизложенное, можно выявить следующий ряд характеристик, которыми должна обладать среда проектирования и вычисления математических функций:

- возможность использования среды как блока проектирования и вычисления математических выражений в других системах;
- обеспечение высокой скорости вычисления;
- возможность динамической генерации кода математических функций;
- обеспечение простоты описания математических функций.

Разработка системы оптимизации вычисления математических функций. Учитывая рассмотренные во введении способы реализации вычислений математических функций и требования к среде проектирования, необходимо для ускорения вычислений в качестве способа реализации использовать описание математических функций на языке ассемблер. Как уже отмечалось, минусом этого подхода является необходимость знания архитектурных и программных особенностей ЭВМ, на которой производятся вычисления. Рассмотрим эти особенности для ЭВМ, построенных на базе процессоров фирмы Intel серии AI-32/64. В этих микропроцессорах реализован математический сопроцессор, позволяющий ускорить вычисление следующих математических функций [1–6]:

- линейные ($f = k \cdot a + b - \frac{y}{c}$);
- циклические ($f = \sum_{i=0}^{100000000} \frac{y \cdot x}{z^a + \frac{z}{x}}, f = \prod_{i=0}^{100} \frac{100}{z^a + \frac{z}{x}}$);
- степенные ($f = x^n$), коренные ($f = x^{\frac{1}{n}}$);

- экспоненциальные ($f = e^x$);
- показательные и логарифмические ($y * \log_2(x+1)$, $\log_2 2e$, $y * \log_2(x)$ и т.п);
- тригонометрические ($\cos(x)$, $\sin(x)$ и т.д.) [7].

Также следует отметить, что при реализации математических функций существует определённый набор трудновычисляемых функций. Трудновычисляемые функции – это функции, в которых зависимость времени вычисления от значения аргумента имеет экспоненциальный характер. Примером такой функции может являться факториал.

Математические модели редко бывают представлены в виде отдельных вышеперечисленных функций, обычно они состоят из некоторой комбинации этих функций. В этом случае необходимо анализировать математическую функцию до составления машинного кода на возможное наличие повторяющихся блоков функции, константных выражений и т.д. Для осуществления анализа и абстрагирования при построении математических функций от особенностей конкретной ЭВМ необходимо ввести промежуточный байт-код.

Таким образом, предлагается структура системы оптимизации вычислений математических функций (СОМФ), изображённая на рисунке 1.

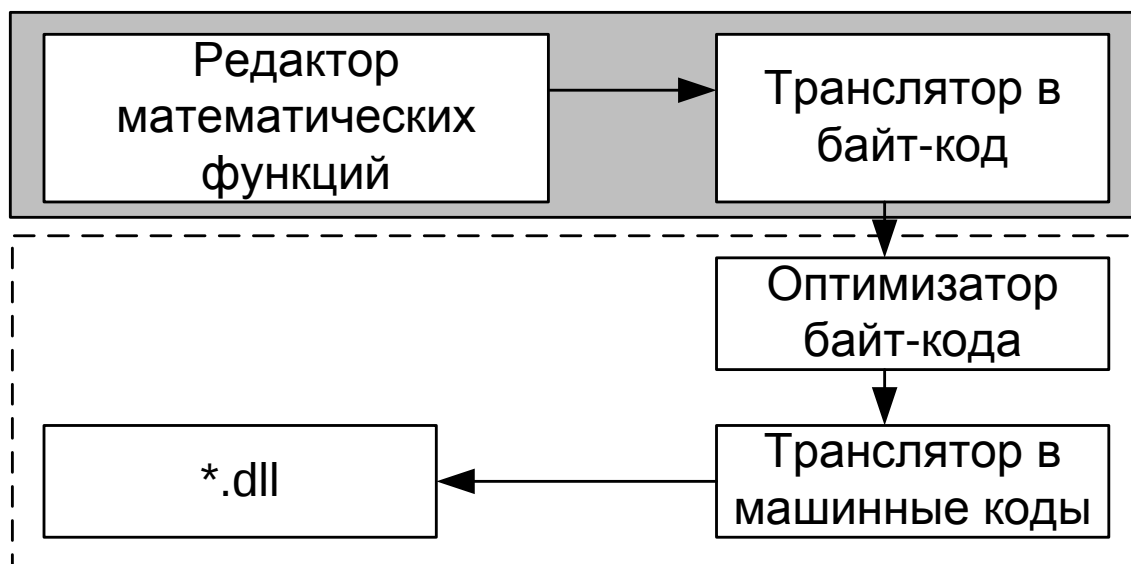


Рис. 1. Структурная схема среды вычисления математических выражений

Структура состоит из четырёх блоков: редактирование математических функций, транслятор в байт-код, оптимизатор байт-кода, транслятор в машинные коды.

Первый и второй блоки представлены в виде отдельной программы, на выходе которой получается файл байт-кода. Такой подход позволяет подключать редакторы математических функций сторонних разработчиков. Третий и четвёртый блоки представлены в виде динамически подключаемой библиотеки. Это позволяет осуществлять встраивание данной системы в другие программные комплексы как составной части.

Код, находящийся в dll-библиотеке, работает только с математическим сопроцессором, который поддерживает достаточное количество форматов данных и команд их обработки.

В сопроцессоре реализованы команды целочисленных и вещественных арифметических операций, вычисления тригонометрических, степенных и других функций, обеспечивающие независимость, гибкость и максимальную скорость, которая в сравнении с центральным процессором превышает его в 100 раз.

На рисунке 2 изображён процесс преобразования математического представления функции в динамическую библиотеку.



Рис. 2. Процесс преобразования математического представления функции в динамическую библиотеку

Экспериментальные исследования. При проведении экспериментальных исследований сравнивались временные затраты вычисления тестовых функций в трёх средах проектирования – MathCad, Microsoft Visual Studio 2005 язык C++ и СОФМ. Для тестирования были выбраны следующие тестовые функции:

– линейная $x = a \cdot b + \frac{c}{d}$;

– циклическая $\sum_{i=0}^{100000000} \frac{y \cdot x}{z^a + \frac{z}{x}}$;

– тригонометрическая $A = \cos(\pi) + \frac{\sin(n!)}{\operatorname{arctg}(\frac{1}{x})}$ и $\sum_{n=1}^{10000000} \frac{3 \cdot n}{2 \cdot n + x!}$.

Тестирование проводилось на аппаратной платформе IBMPC с центральным процессором семейства IntelPentiumIV 1,616 ГГц.

Результаты проведения экспериментальных исследований приведены в таблице и на рисунке 3.

Результаты экспериментов

Функция	Mathcad	C++	СОМФ		
	время вычисления t , сек	время вычисления t , сек	время вычисления t , сек	прирост скорости вычисления относительно Mathcad	прирост скорости вычисления относительно C++
Линейная	0	0	0	0	0
Циклическая	0,7	0,4	0,15	4,67	2,67
Трудновычисляемая	34	4,3	0,5	68	8,60
Тригонометрическая	219	24,2	0,6	365	40,33

На рисунке 3 приведена зависимость времени вычисления тестовых функций от среды проектирования. Из графика и таблицы видно, что время вычисления линейной функции во всех системах одинаково, а при вычислении тригонометрической функции СОМФ – в 4,67 раз быстрее, чем MathCad, и в 2,67 раза быстрее реализации функции на языке C++.

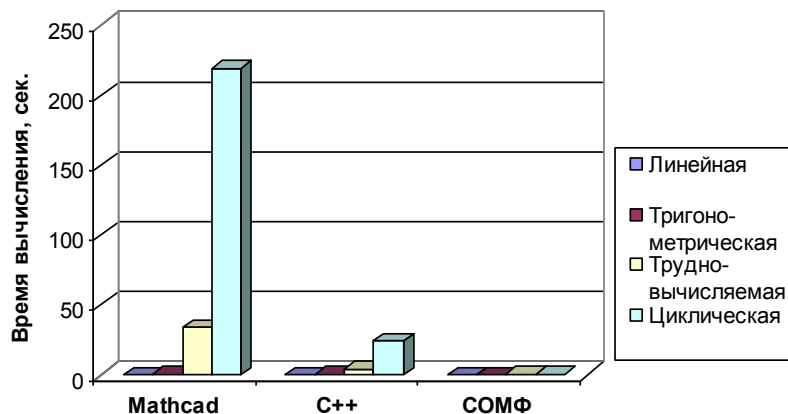


Рис. 3. Время вычисления тестовых функций

Время, затраченное на вычисление трудновычисляемой функции и тригонометрической в среде СОМФ, в 68 и 365 раз быстрее, чем в MathCad, и в 8,60 и 40,33 раз быстрее реализации на языке С++ соответственно.

Заключение. Экспериментальные исследования показали правильность выбранного подхода для проектирования и вычисления математических функций на ЭВМ. Разработанная система обладает широкими возможностями моделирования высокоскоростных вычислений на аппаратном уровне. Результаты экспериментальных исследований показывают многократное уменьшение времени расчёта сконструированных математических функций при помощи СОМФ.

Также следует отметить, что СОМФ можно улучшить за счёт введения поддержки таких технологий, как MMX и SSE, позволяющих ускорить операции работы с последовательностями данных (вектор, матрица и т.д.).

Список литературы

1. Пирогов В.Ю. Ассемблер для Windows. – 3-е изд., перераб. и доп. – СПб.: БХВ – Петербург, 2005. – 864 с.: ил.
2. Юров В.И. Assembler: учебник для вузов. – 2-е изд. – СПб.: Питер, 2005. – 637 с.: ил.
3. Магда Ю.С. Ассемблер. Разработка и оптимизация Windows-приложений. – БХВ-Петербург, 2003. – 544 с.: ил.
4. Ахо А., Сети Р., Ульман Дж. Компиляторы: принципы, технологии и инструменты: пер. с англ. – М.: Вильямс, 2003. – 768 с.: ил.
5. Магда Ю.С. Использование ассемблера для оптимизации программ на С++. – СПб.: БХВ – Петербург, 2004. – 496 с.: ил.
6. Магда Ю.С. Ассемблер для процессоров Intel Pentium. – СПб.: Питер, 2006. – 410 с.: ил.
7. Кремер Н.Ш., Путко Б.А., Тришин И.М., Фридман М.Н. Высшая математика для экономистов: учебник для вузов / под. ред. проф. Н.Ш. Кремера. – 2-е изд., перераб. и доп. – М.: ЮНИТИ, 2002. – 471 с.
8. Черняк А.И., Новиков В.И., Мельников О.И., Кузнецов А.И. Математика для экономистов на базе MathCad. – СПб.: Петербург, 2003. – 496 с.: ил.

РАЗДЕЛ 5

МОДЕЛИРОВАНИЕ ТЕХНИЧЕСКИХ СИСТЕМ

УДК 681.3:519

А.А. Токарев, В.Е. Мешков

ФОРМАЛИЗАЦИЯ ПРОЦЕДУРЫ СВЁРТКИ ТОПОЛОГИЧЕСКИ СЛОЖНЫХ УЧАСТКОВ ГИДРАВЛИЧЕСКИХ СИСТЕМ ПАРОВЫХ КОТЛОВ НА ОСНОВЕ СВОЙСТВ МАТРИЦЫ ИНЦИДЕНТНОСТИ

В данной статье изложены принципы формализации процедур свёртки топологически сложных участков гидравлических систем паровых котлов. Формализация основана на использовании свойств матрицы инцидентности графа исходного описания гидравлической системы котла. На основе свойства матрицы инцидентности разработан алгоритм для свёртывания расчётной схемы.

Ключевые слова: формализация, процедуры свёртки, топологическая сложность, гидравлическая система, паровой котёл, свойства матрицы инцидентности.

Введение. В соответствии с системным подходом к моделированию систем первоначально разрабатывается топологическая математическая модель, а далее – математические модели компонентов моделируемой системы [3].

Гидравлические системы котельных агрегатов имеют большое количество компонентов, соединённых как последовательно, так и параллельно. При этом число компонентов в несколько раз превышает количество контуров [1]. При записи топологических уравнений гидравлических систем паровых котлов удобно использовать промежуточную графическую форму – представление модели в виде эквивалентной схемы, состоящей из двухполюсных элементов. Общность подхода при этом сохраняется, так как любой многополюсный компонент можно заменить подсхемой из двухполюсников [5].

В последние годы для анализа ряда сложных систем широко и успешно применяется теория графов. Графы широко используются как структурные модели физических систем, допускающих идеализированное представление в виде схем с сосредоточенными параметрами. Графом $V = (im, jm)$, в общем случае [4], называется множество вершин (узлов) im ,

связи между которыми определены множеством рёбер jm . Если рёбра ориентированы, что показывается стрелкой, то они называются дугами, и граф с такими рёбрами называется ориентированным графом или орграфом. Вследствие этого эквивалентную схему можно рассматривать как ориентированный граф, дуги которого соответствуют ветвям схемы. Направления потоков в ветвях выбираются произвольно (если реальное направление при моделировании окажется противоположным, то это приведёт лишь к отрицательным численным значениям потока).

Исходное описание графа гидравлической системы котла

Исходное описание графа гидравлической системы котла даёт его матрица соединений (инцидентности) $[a_{ij}]$, число горизонтальных строк которой равно числу узлов в схеме, а число столбцов – числу дуг. Элементы этой матрицы $[a_{ij}]$ определяются следующим образом:

$$[a_{ij}] = \begin{cases} -1, & \text{если } j\text{-я дуга входит в } i\text{-й узел;} \\ 1, & \text{если } j\text{-я дуга выходит из } i\text{-го узла;} \\ 0, & \text{если } j\text{-я дуга и } i\text{-й узел не инцидентны.} \end{cases}$$

Обычно в расчётных схемах около шести и более компонентов соединены последовательно [1]. В соответствии с уравнением неразрывности [5] при последовательном соединении расход в компонентах одинаков, однако при решении системы нелинейных уравнений расход определяется в каждом компоненте независимо от соединения, что приводит к неоправданно большому количеству уравнений. Чтобы снизить порядок системы уравнений и тем самым повысить вычислительную эффективность математической модели системы, желательно выполнить предварительное преобразование модели перед её многошаговым численным решением. Предварительное преобразование сводится к исключению из системы части неизвестных и соответствующего числа уравнений.

Алгоритм для свёртывания расчётной схемы

Для уменьшения количества уравнений необходимо произвести замену последовательно соединённых компонентов одним эквивалентным компонентом, т.е. произвести свёртку схемы.

На основе свойства матрицы инцидентности разработан алгоритм для свёртывания расчётной схемы, основные этапы которого представлены ниже:

- по известному графу расчётной схемы строим матрицу инцидентности $[a_{ij}]$, при этом номер строки матрицы соответствует номеру узла на расчётной схеме;

- последовательно просматриваем строки матрицы $[a_{ij}]$. Если узел не моделирует «вход» или «выход», тогда суммируем элементы строки (S_{Σ}) и находим количество ненулевых элементов (K). Таким образом, осуществляется поиск последовательно соединённых элементов;
- если S_{Σ} равно нулю и K равно двум, тогда столбец, содержащий в рассматриваемой строке матрицы $[a_{ij}]$ минус единицу, суммируется со столбцом, содержащим в рассматриваемой строке единицу. Это действие означает замену двух последовательно соединённых компонентов одним, при этом узел, разделяющий эти компоненты, удаляется.

Рассмотрим работу данного алгоритма на примере свёртывания графа расчётной схемы, который представлен на рисунке 1. Количество элементов графа равно восьми, количество узлов – семи. При этом первый узел моделирует «вход», а четвёртый и седьмой узлы – «выход».

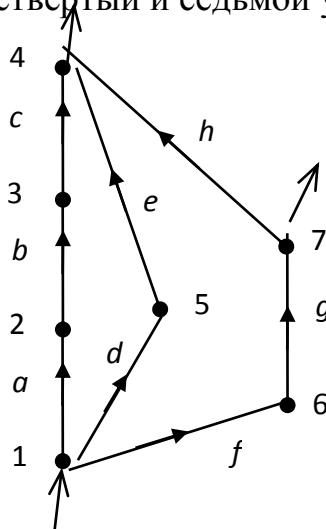


Рис. 1. Граф расчётной схемы

Матрица инцидентности данного графа будет выглядеть следующим образом:

$$[a_{ij}] = \begin{matrix} & \begin{matrix} a & b & c & d & e & f & g & h \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix} \end{matrix}.$$

Так как первый узел моделирует «вход», то переходим сразу к рассмотрению второй строки матрицы $[a_{ij}]$. Для этой строки S_Σ равна нулю и K равно двум. После суммирования столбцов a и b и удаления второй строки, содержащей нули, имеем:

$$\begin{array}{c} \\ 1 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} a+b \\ c \\ d \\ e \\ f \\ g \\ h \end{array} \left(\begin{array}{ccccccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{array} \right).$$

В третьей строке, как и во второй, S_Σ равна нулю и K равно двум, поэтому производим суммирование столбцов $a+b$ и c . После суммирования третью строку удаляем. В результате проделанных операций матрица $[a_{ij}]$ будет выглядеть следующим образом:

$$\begin{array}{c} \\ 1 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} a+b+c \\ d \\ e \\ f \\ g \\ h \end{array} \left(\begin{array}{cccccc} 1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array} \right).$$

Четвёртый узел моделирует «выход», поэтому четвёртую строку матрицы пропускаем и переходим к следующей строке. Для пятой строки сумма элементов равна нулю и количество ненулевых элементов равно двум, поэтому суммируем столбцы d и e , а строку пять удаляем. В результате имеем следующую матрицу:

$$\begin{array}{c} \\ 1 \\ 4 \\ 6 \\ 7 \end{array} \begin{array}{c} a+b+c \\ d+e \\ f \\ g \\ h \end{array} \left(\begin{array}{ccccc} 1 & 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{array} \right).$$

В шестой строке выполняются те же условия, что и для удаляемых ранее строк. Поэтому её также удаляем, предварительно просуммировав столбцы f и g . Последняя, седьмая, строка, как и четвёртая, моделирует «выход». Эту строку пропускаем, и она была последней в матрице инцидентности. Граф свёрнутой расчётной схемы представлен на рисунке 2, а его матрица инцидентности будет выглядеть следующим образом:

$$\begin{array}{c} a+b+c \quad d+e \quad f+g \quad h \\ \begin{array}{c} 1 \\ 4 \\ 7 \end{array} \left(\begin{array}{cccc} 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{array} \right).$$

До свёртывания расчётной схемы для определения расходов в элементах необходимо было решить систему из восьми уравнений. После свёртки размер решаемой системы уменьшается до четырёх.

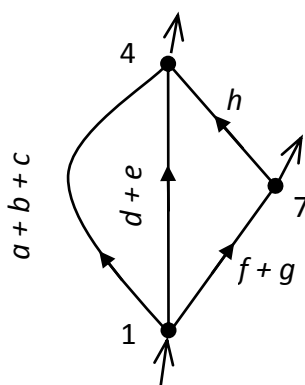


Рис. 2. Граф свёрнутой расчётной схемы

Таким образом, свёртывание расчётной схемы позволяет в несколько раз уменьшить размер решаемой системы уравнений, что приводит к уменьшению времени счёта, а также к повышению устойчивости решения системы.

Заключение

В рассмотренном подходе повышение устойчивости решения происходит вследствие того, что гидравлическая характеристика компонентов, которые заменили последовательно соединённые компоненты, выравнивается.

Список литературы

1. Клер А.М., Скрипкин С.К., Деканова Н.П. Автоматизация построения статических и динамических моделей теплоэнергетических установок // Изв. РАН. Энергетика. – 1996. – № 3. – С. 78-84.
2. Корн Г., Корн Т. Справочник по математике (для научных работников и инженеров). – М.: Наука, 1973. – 832 с.
3. Норенков И.П. Основы автоматизированного проектирования: учебник для вузов. – М.: МГТУ им. Н.Э. Баумана, 2000. – 360 с.
4. Сигорский В.П. Математический аппарат инженера. – 2-е изд., стер. – Киев: Техника, 1977. – 768 с.
5. Попырин Л.С., Самусев В.И., Эпельштейн В.В. Автоматизация математического моделирования теплоэнергетических установок. – М.: Наука, 1981. – 236 с.

УДК 681.5

В.В. Семёнов, Ю.Б. Ханжонков, Ю.Г. Асцатуров

ОБРАЗЦОВЫЙ ТЕЛЕВИЗИОННЫЙ АНАЛИЗАТОР ПРОМЫШЛЕННЫХ АЭРОЗОЛЕЙ

В работе представлен образцовый телевизионный анализатор промышленных аэрозолей, позволяющий получать в трёх проекциях основные геометрические и оптические параметры промышленного аэрозоля, необходимые для градуировки простых оптико-электронных датчиков. Приведён пример алгоритма распознавания с помощью системы LABView, позволяющий в режиме реального времени получать данные о форме, размерах и цвете исследуемых промышленных аэрозолей.

Ключевые слова: оптико-электронные датчики, телевизионный анализ, лазерный фотоэлектрический счётчик аэрозолей, распознавание промышленного аэрозоля, цифровая обработка изображений, NI Vision Builder, LabVIEW.

При разработке оптико-электронных датчиков для контроля концентрации промышленного аэрозоля существует проблема их градуировки под конкретный вид пыли. В ряде случаев приходится градуировать вышеупомянутые датчики только на месте их установки или в лабораторных условиях на специальных стендах.

Используя предлагаемый телевизионный анализатор, позволяющий определять и хранить параметры для расчёта и градуировки оптико-электронных датчиков, можно достаточно быстро перенастраивать существующие датчики под конкретный вид промышленных аэрозолей. Это позволяет сократить сроки разработки оптико-электронных датчиков и расширить диапазон их применения [1].

Описанный ниже телевизионный анализатор формирует базу данных, характеризующих форму и оптические параметры промышленных аэрозолей, и может быть использован в качестве эталонного средства при градуировке оптико-электронных датчиков, основанных на ослаблении и рассеянии оптического излучения промышленным аэрозолем.

Среди различных методов определения размеров и формы взвешенных частиц наиболее информативными являются телевизионные методы [2]. Обычно регистрируется одно изображение каждой частицы, соответствующее её проекции на плоскость, перпендикулярную оптической оси светового пучка [3]. Существенно повысить информативность измерений данным методом позволяет получение двух и более (для более сложных форм) изображений частиц.

На рисунке 1 приведена схема, позволяющая получить в плоскости регистрации (на ПЗС видеокамеры) одновременно три изображения каждой частицы, соответствующие её проекциям на три взаимно перпендикулярные плоскости.

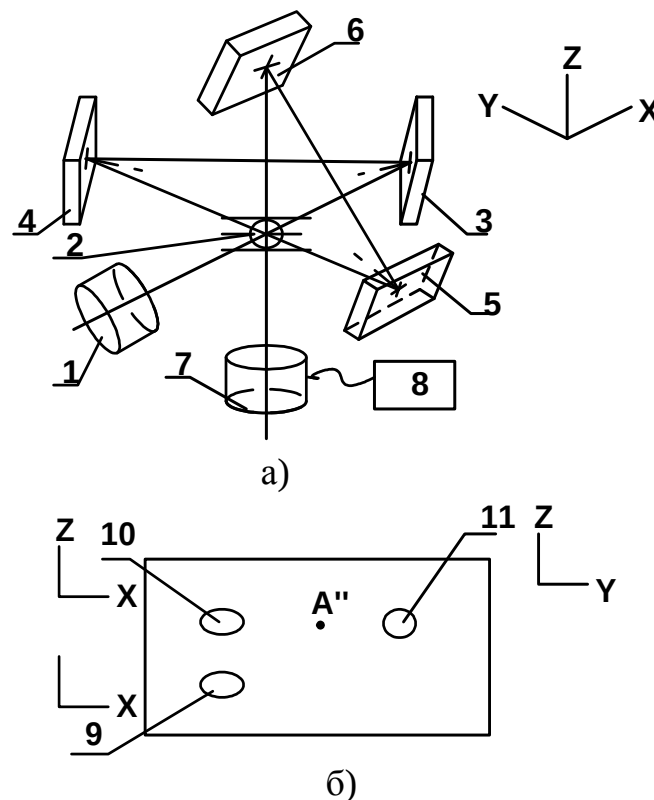


Рис. 1. Формирование проекций частицы на три взаимно перпендикулярные плоскости:

а – общая схема; б – вид плоскости регистрации;

1 – осветитель; 2 – поток частиц; 3 – 6 – зеркала; 7 – видеокамера;

8 – блок обработки; 9 – третье изображение частицы;

10 – отображение первого изображения; 11 – отображение второго изображения

Из рисунке 1 видно, что в плоскости регистрации имеется три проекции частицы. При этом, поскольку частица смещена относительно осей OX , OY на расстояния, превосходящие соответствующие размеры частиц, то все три изображения пространственно разнесены.

Функциональная схема разработанного лабораторного анализатора, использующего разработанные оптические схемы, показана на рисунке 2.

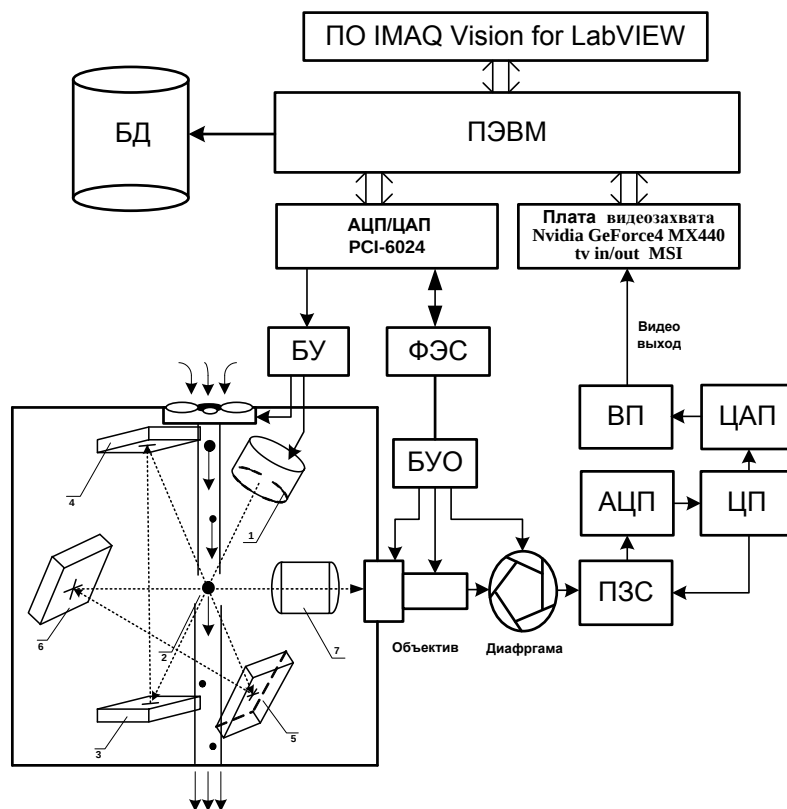


Рис. 2. Функциональная схема разработанного телевизионного анализатора:

1 – лазерный осветитель; 2 – поток частиц; 3–6 – зеркала; 7 – микролинза;

БУ – блок управления вентилятором и лазерным излучателем;

ФЭС – фотоэлектрический счётчик (ФЭС-1); БУО – блок управления объективом (моторами фокусировки, трансфокатора и диафрагмы) видеокамеры; ПЗС – видео-сенсор (прибор с зарядовой связью); АЦП – аналогово-цифровой преобразователь;

ЦП – центральный процессор; ЦАП – цифро-аналоговый преобразователь;

ВП – видеопроцессор; БД – база данных параметров частиц

Динамический ввод аэрозолей в область контроля осуществляется прокачной кюветой с шириной канала 1,5 мм и вакуумным насосом. Скорость прокачки подбирается программно ПЭВМ так, чтобы в каждый кадр видеосигнала попадало не более 1 частицы (50 частиц в секунду), для чего в функциональную схему телевизионного анализатора введён фотоэлектрический счётчик частиц пыли (ФЭС), сопряжённый через плату ввода вывода NI PCI-6024 с ПЭВМ. Программа LAbVIEW проводит амплитудный и частотный анализ входного сигнала с ФЭУ ФЭС с целью оператив-

ного (предварительного) выявления размеров частиц. Сформированный программой сигнал управления поступает с ПЭВМ и в плату ввода – вывода NI PCI-6024 для управления скоростью прокачки (вакуумным насосом) и фокусировкой объектива видеокамеры телевизионного анализатора.

На рисунке 3 представлена функциональная схема лазерного счётчика частиц.

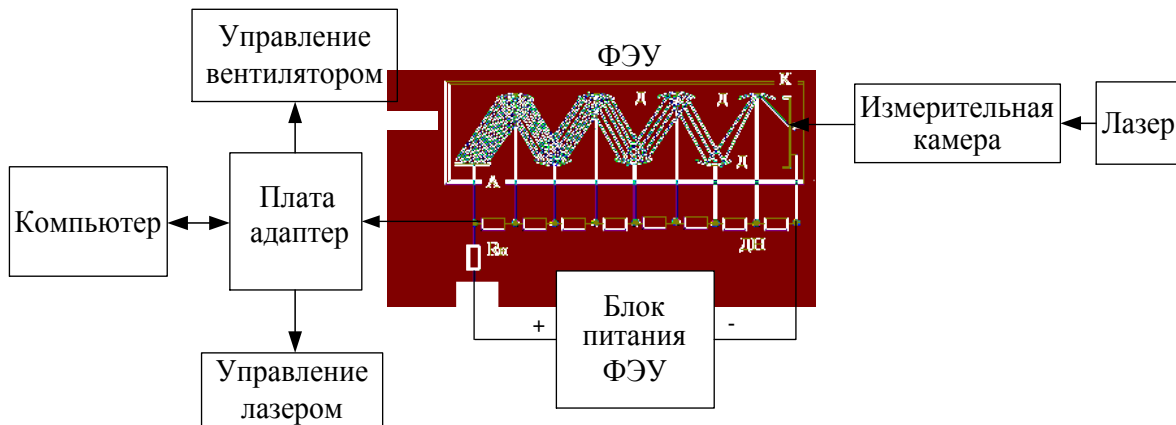


Рис. 3. Функциональная схема лазерного ФЭС

Частицы пыли в пробе воздуха заставляют лазерный луч создавать рассеянный свет в измерительной камере устройства. Луч, сфокусированный в пучок системой линз, улавливается фотоэлектронным умножителем (ФЭУ-68) и преобразуется в электрический сигнал. Далее сигнал поступает на плату сбора данных (NI PCI-602), сопряжённую с ПЭВМ. Результат измерения отображается на дисплее компьютера. Управление лазером и вентилятором производится программно, с использованием управляющих транзисторов.

На рисунке 4 представлена разработанная виртуальная панель ФЭС в оболочке labview.

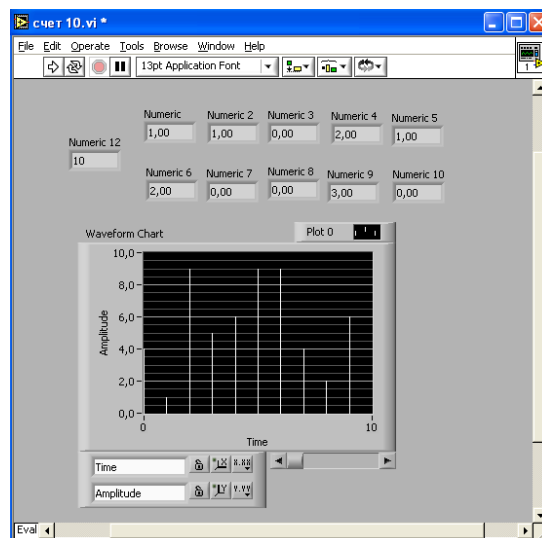


Рис. 4. Внешний вид виртуального прибора лазерного ФЭС

Внешний вид образцового анализатора промышленных аэрозолей показан на рисунке 5.

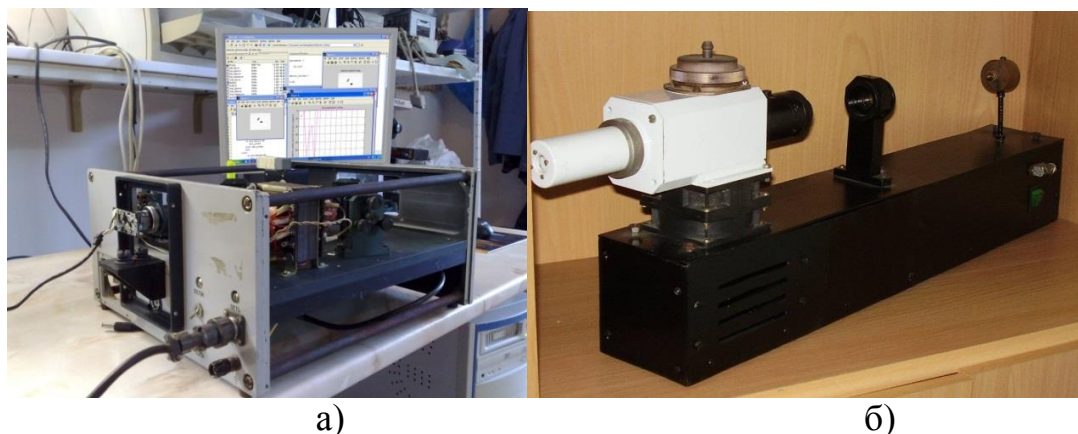


Рис. 5. Внешний вид образцового телевизионного анализатора промышленных аэрозолей: (а) телевизионный анализатор (б) лазерного ФЭС

На рисунке 6 показаны изображения, полученные образцовым телевизионным анализатором.

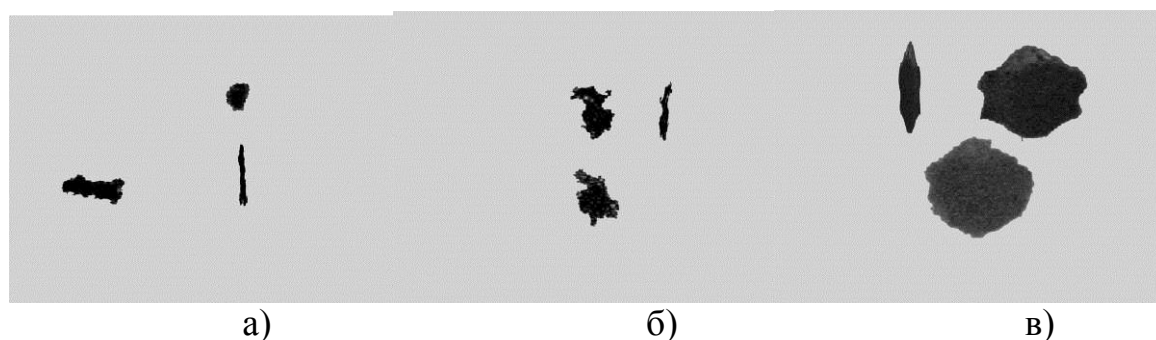


Рис. 6. Изображения аэрозольных частиц различной формы: а) 10 мкм; б) 30 мкм; в) 100 мкм (каждый кадр (частица в трёх проекциях) соответствует одной частице)

Следует отметить, что в предлагаемом анализаторе нижний предел измерений по размерам (определяемый разрешающей способностью формирующей оптики) составляет примерно 1 мкм, а верхний предел измерений по концентрации может составлять 10^5 см^3 .

Рассмотрим алгоритм функционирования системы телевизионного анализа изображений.

Наблюдаемые частицы могут иметь самые различные формы [4]:

- изометрические частицы, для которых в первом приближении все три размера совпадают. К этому классу относятся сферы, правильные многогранники или частицы, близкие к ним по форме. Большинство научных представлений о поведении аэрозолей относится к изометрическим частицам;

- пластинки – частицы, имеющие два длинных размера и один короткий, к этому классу относятся лепестки или их кусочки, чешуйки, диски;
- волокна – частицы, протяжённые в одном направлении и имеющие сравнительно небольшие размеры в двух других; к их числу относятся призмы, иглы, нити или минеральные волокна, например волокна асбеста.

В зависимости от различных размеров и форм частиц будут использоваться применяемые методы цифровой обработки и их параметры.

Составим алгоритм (рис. 8), основываясь на выделении обработок по предварительному анализу форм и размеров частиц.

Для начала нужно определить форму частицы. Если все три размера совпадают – это изометрическая (сферическая) частица, если нет – значит, частица – либо пластина, либо волокно.

Далее каждый из классов разделим ещё на две группы, распределив по размерам.

С каждой из групп проведём следующие операции (рис. 7, 8) [5]:

- считывание и отображение изображений;
- улучшение контраста изображений, сглаживание;
- пространственная (линейная) фильтрация: использование медианного фильтра для удаления шума на изображении;
- бинаризация изображений, пороговое разделение, выделение кромок;
- морфологический анализ;
- количественный анализ;
- построение гистограммы распределения по размерам аэрозолей;
- вычисление данных, характеризующих форму, размер и цвет аэрозоля, сброс информации на диск.

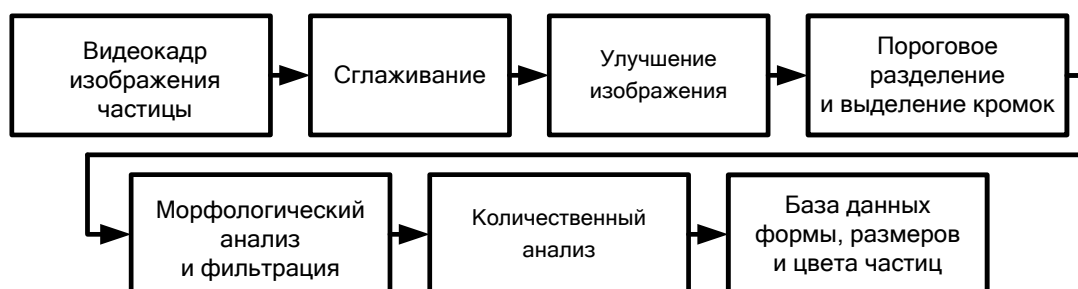


Рис. 7. Процесс распознавания размера и формы аэрозолей

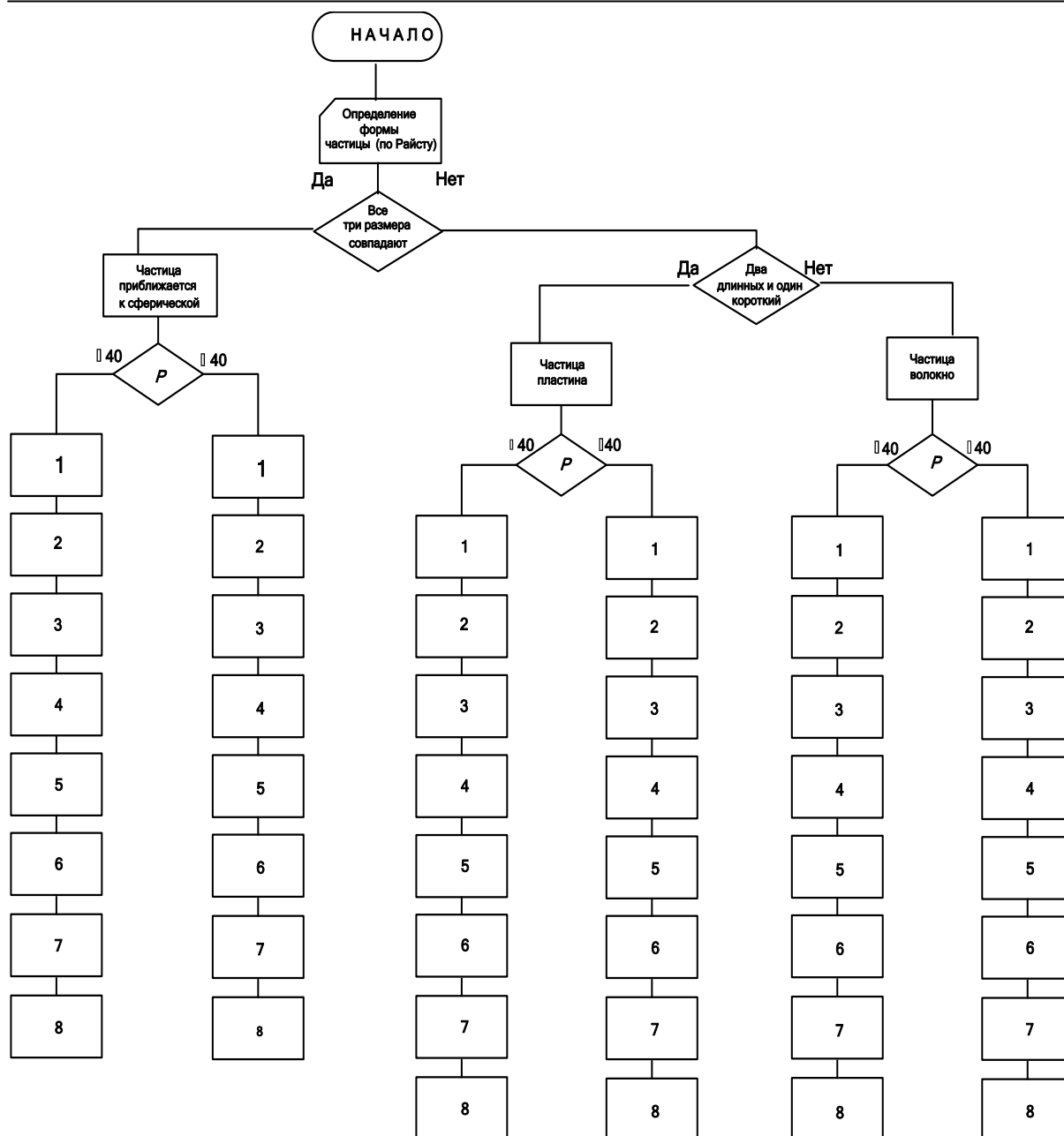


Рис. 8. Алгоритм функционирования телевизионного анализатора

Обзор существующего программного обеспечения показал, что наиболее перспективными программными продуктами для телевизионного анализа изображений являются программы MATLAB и NI Vision Builder.

На рисунке 9 приведён результат работы программы NI Vision Builder по вышеприведённому алгоритму работы.

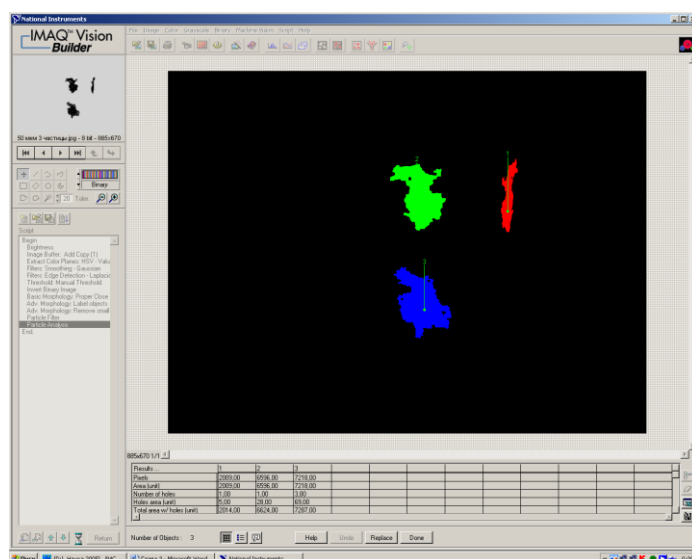


Рис. 9. Результаты работы программы анализа формы и размеров аэрозольных частиц (размер 30 мкм)

Полученные данные хранятся в базе данных или в таблице Excel.

На рисунке 10 показан результат компилирования, написанного в NI Vision Builder скрипта в оболочку LabVIEW.

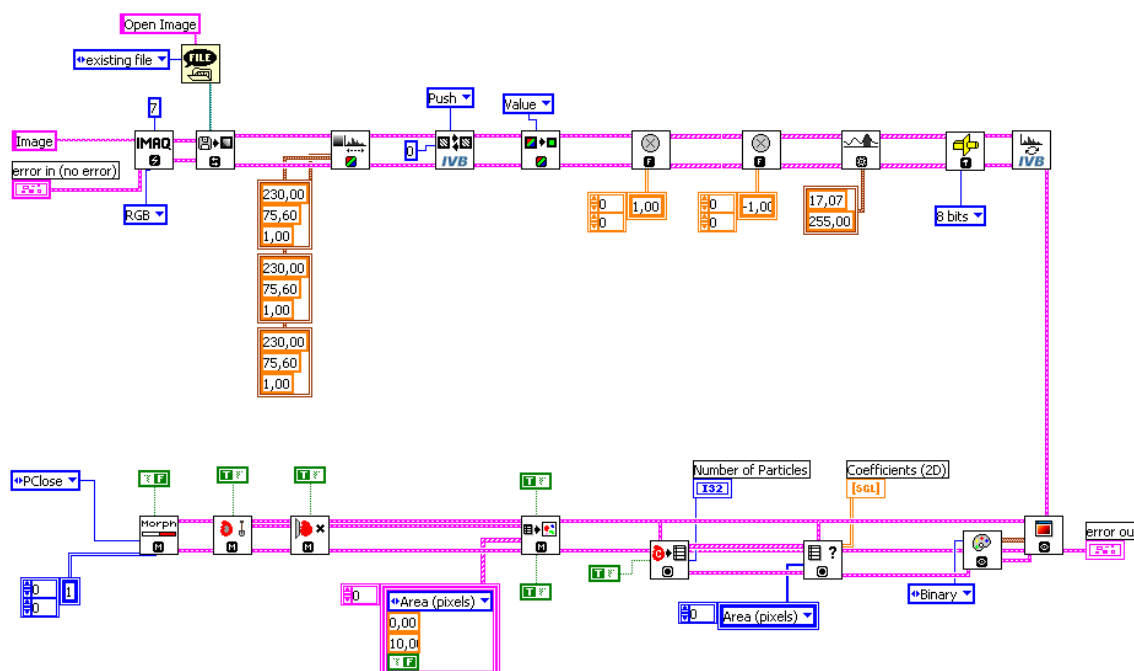


Рис. 10. Блок-диаграмма обработки изображений частиц пыли для образцового телевизионного анализатора

В настоящее время авторами работы проводится усовершенствование алгоритмов обработки анализатора и разработка специализированного программного обеспечения для моделирования и разработки оптико-электронных датчиков, использующего получаемую обширную базу данных о размерах, форме и цвете аэрозолей (около 60 параметров для одной частицы). В основе моделирования лежат теория Ми и её клоны для частиц различных форм [6].

Таким образом, предлагаемый в работе анализатор позволяет определять и хранить параметры промышленных аэрозолей, необходимые для моделирования, расчёта и градуировки оптико-электронных датчиков контроля аэрозолей (запылённости). Используя полученные при помощи анализатора данные, можно достаточно быстро перенастраивать существующие датчики под конкретный вид промышленных аэрозолей. Это позволяет сократить сроки разработки оптико-электронных датчиков и расширить диапазон их применения.

Список литературы

1. Семёнов В.В. Оптико-электронный пылемер системы автоматического управления проветриванием шахт: дис. ... канд. техн. наук. – Новочеркасск, 2002. – 142 с.
2. Оптико-электронные методы изучения аэрозолей / С.П. Беляев, Н.К. Никифорова, В.В. Смирнов [и др.]. – М.: Энергоиздат, 1981. – 232 с.
3. Коломиец Г.А., Коломиец С.М. Анализатор размеров и формы аэрозолей «АРФА» // Оптика атмосферы и океана. – 1999. – Т. 12, № 6. – С. 553–555.
4. Райст П. Аэрозоли. Введение в теорию: пер. с англ. – М.: Мир, 1987. – 200 с.
5. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
6. Борен К., Хафмен Д. Поглощение и рассеяние света малыми частицами: пер. с англ. – М.: Мир, 1986. – 660 с.

УДК 544.643

Н.Е. Галушкин, Н.Н. Галушкина

СТРУКТУРНОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССА САМОРАЗРЯДА В ЩЕЛОЧНЫХ АККУМУЛЯТОРАХ

Предложена структурная модель щелочного аккумулятора с учётом саморазряда, из которой получены все наиболее известные эмпирические зависимости. Показано, что существующие эмпирические зависимости, описывающие процесс саморазряда в щелочных аккумуляторах, не противоречат, а дополняют друг друга, так как они справедливы каждое в своём интервале рассмотрения времени саморазряда.

Ключевые слова: структурное моделирование, щелочные аккумуляторы, само-разряд.

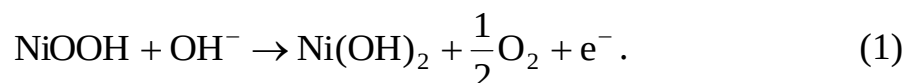
Введение. В теории импеданса довольно широко используется структурное моделирование при исследовании различных электрохимических процессов и систем [1].

Структурное моделирование основывается на системном подходе, при котором исследуемый объект рассматривается как система, состоящая из подсистем или элементов [2]. Отдельные модельные компоненты находятся в непосредственной близости, не проникая при этом один в другой. Взаимодействие между ними осуществляется через разделяющие их поверхности и связи.

Тем самым структурная модель электрохимической системы представляет собой некоторую электротехническую схему. Структурные импедансные модели строятся из элементов, которые по своему физическому смыслу соответствуют моделируемым процессам. При этом элементами структурной модели могут быть как стандартные электротехнические элементы, так и специфические электрохимические элементы. Структурные модели всегда очень наглядны и понятны как с электротехнической точки зрения, так и с электрохимической, что является их бесспорным преимуществом. Однако до недавнего времени структурное моделирование использовалось только в теории импеданса, то есть в моделях с малыми токами. В работе [3] показано, что методы структурного моделирования могут быть с успехом применены и при моделировании процессов разряда в аккумуляторах при больших рабочих токах.

В данной статье исследуется процесс саморазряда в щелочных аккумуляторах на базе структурной модели аккумулятора с саморазрядом.

Основная электрохимическая причина саморазряда никель-кадмиевых (НК) аккумуляторов связана с тем, что потенциал оксидно-никелевого электрода (ОНЭ) положительнее потенциала обратимого кислородного электрода, поэтому на ОНЭ может идти реакция разряда гидроксил ионов с выделением газообразного кислорода, сопровождающаяся восстановлением никеля [4]:



При протекании описанной реакции снижается потенциал оксидно-никелевого электрода и теряется ёмкость, эквивалентная выделившемуся количеству кислорода. По мере снижения потенциала ОНЭ замедляется разряд ионов гидроксила с образованием газообразного кислорода. В этом случае основной причиной саморазряда ОНЭ становится протекание реакции непосредственного взаимодействия гидроксидов металлов, содержащих активный кислород, с прилегающими поверхностями графита или металлокерамической основы [5, 6]. Саморазряд кадмиевого электрода связан в основном с химическим окислением кадмия кислородом, выделившимся на ОНЭ [7]. Конечно, существуют и другие причины саморазряда [8, 9] и т.д.

Структурная модель саморазряда щелочных аккумуляторов
 Простейшая структурная модель аккумулятора с учётом саморазряда будет иметь вид (рис. 1):

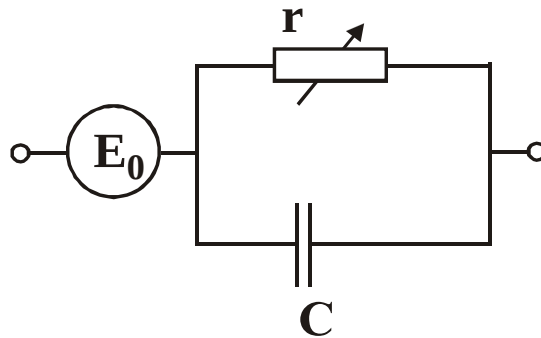


Рис. 1. Простейшая структурная модель щелочного аккумулятора с учётом саморазряда:

E_0 – идеальный конденсатор постоянного напряжения, моделирующий ЭДС аккумулятора после очень большого срока хранения (теоретически бесконечного);
 C – псевдоконденсатор, моделирующий процесс саморазряда аккумулятора, т.е. изменение напряжения на его обкладках; r – нелинейное сопротивление, моделирующее электрохимические процессы саморазряда на границе активного вещества и электролита

Так как ток саморазряда обусловлен необратимыми электрохимическими реакциями (1) и др., то он в общем случае должен описываться сильно несимметричной функцией замедленного разряда, то есть функцией вида [10]:

$$i_c(u) = i_0 \left\{ \exp\left(\frac{\alpha z F u}{RT}\right) - \exp\left(-\frac{(1-\alpha) z F u}{RT}\right) \right\}, \text{ при } \alpha \approx 1, \quad (2)$$

где i_0 – ток обмена; F – число Фарадея; R – газовая постоянная; T – абсолютная температура; u – напряжение на нелинейном сопротивлении r .

В этом случае саморазряд псевдоконденсатора C (рис. 1) будет описываться уравнением:

$$C \frac{du}{dt} + i_c(u) = 0, \quad (3)$$

где $i_c(u)$ – ток утечки (саморазряда), через нелинейное сопротивление r .

Начальное условие для уравнения (3) будет:

$$u|_{t=0} = u_0 = E_0 - E_{\Pi}, \quad (4)$$

где E_0 – ЭДС заряженного аккумулятора; E_{Π} – предельная ЭДС, т.е. ЭДС, до которой изменяется напряжение на клеммах аккумулятора при теоретически бесконечном сроке хранения в соответствии с конкретным механизмом саморазряда. Например, при саморазряде в соответствии с электрохимической реакцией (1) предельная ЭДС E_{Π} будет определяться потенциалом обратимого кислородного электрода.

Модель 1. Пусть ток саморазряда описывается уравнением Тафеля:

$$i_c(u) = i_0 \exp(au), \quad (5)$$

где $a = \frac{\alpha z F}{RT}$.

Формула (5) может быть использована вместо формулы (2) при достаточно больших u , т.е. в достаточно широком интервале изменения t от начала саморазряда. Решим уравнение (3) при граничных условиях (4) и токе саморазряда (5). В этом случае напряжение на клеммах аккумулятора будет изменяться со временем согласно уравнению:

$$u_k = E_0 - \frac{1}{a} \ln \left[\frac{ai_0 \exp(au_0)}{C} t + 1 \right]. \quad (6)$$

При

$$t > \frac{C}{i_0 \exp(au_0)a} \quad (7)$$

можно пренебречь единицей в квадратных скобках формулы (6), и она переходит в эмпирическую формулу Гинделиса из работы [11]:

$$u = A - B \cdot \ln t, \quad (8)$$

где A, B – константы; причём B – некоторая функция от температуры.

Чтобы получить выражение для потери ёмкости, при саморазряде проинтегрируем ток саморазряда (5) по времени с учётом того, что напряжение на псевдоконденсаторе C равно $u = (u_k - E_n)$ и (6). Получим:

$$q = \frac{C}{a} \ln \left[\frac{i_0 \exp(au_0)a}{C} t + 1 \right]. \quad (9)$$

Данная функция на широком интервале изменения t может быть очень точно аппроксимирована степенной функцией [12], т.е.:

$$q = [i_0 \exp(au_0)] t^n \quad (10)$$

при $0 < n \leq 1$. Причём n будет зависеть от интервала аппроксимации. При небольшом интервале аппроксимации вблизи нуля $n \approx 1$ (разложение в ряд Тейлора). Для интервалов изменения времени, используемых в работе [11], $n < 1$, так как время в данной работе достаточно большая величина в этом интервале. Функция (9) стремится к бесконечности медленнее, чем t , следовательно, n должно быть меньше единицы, что и даёт аппроксимация.

Формула (10) совпадает с эмпирической формулой

$$q = kt^n \quad (11)$$

из работы [11] при

$$k = i_0 \exp(au_0). \quad (12)$$

Если учесть, что ток обмена возрастает с ростом температуры по закону [13]:

$$i_0 = D \exp\left(-\frac{\alpha z F E}{RT} h\right), \quad (13)$$

где E – ЭДС аккумулятора; h – некоторая константа, учитывающая особенности конкретной электрохимической реакции, в частности строение двойного электрического слоя, то получим зависимость k от температуры в виде:

$$\ln k = \ln D - \left(-\frac{\alpha z F (hE + E_{\Pi} - E_0)}{R}\right) \frac{1}{T}. \quad (14)$$

Данная формула совпадает с эмпирической формулой:

$$\ln k = A - \frac{b}{T} \quad (15)$$

из работы [14] при

$$A = \frac{\ln D}{\ln 10}, \quad b = \frac{\alpha z F (hE + E_{\Pi} - E_0)}{R \ln 10}. \quad (16)$$

Модель 2. Решим уравнение (3) для линейной функции саморазряда вида:

$$i_C(u) = \frac{1}{r} u. \quad (17)$$

Данная функция получается из общей функции (2) разложением в ряд Тейлора при

$$|au| < 1, \quad (18)$$

где $r = \frac{1}{i_0 a}$, $a = \frac{zF}{RT}$.

Решим уравнение (3) при граничных условиях (4) и токе саморазряда (17). В этом случае напряжение на клеммах аккумулятора будет изменяться в зависимости от времени хранения по закону:

$$u_k = u_0 e^{-\frac{t}{Cr}} + E_{\Pi}. \quad (19)$$

Аналогично (9) интегрируя ток саморазряда (17) с учётом $u=(u_k-E_{\Pi})$ и (19), получим выражение для потери ёмкости при саморазряде:

$$q = \Delta q_0 (1 - e^{-\frac{t}{Cr}})$$

и для изменения остаточной ёмкости от времени хранения:

$$q_{oc} = \Delta q_0 e^{-\frac{t}{Cr}} + q_{\Pi}, \quad (20)$$

где $\Delta q_0 = C \cdot u_0$ – максимальная потеря ёмкости при саморазряде; $q_n = Q - \Delta q_0$ – предельная остаточная ёмкость аккумулятора, т.е. остаточная ёмкость аккумулятора после бесконечного времени хранения; Q – ёмкость аккумулятора в начале саморазряда.

Формула (20) совпадает с эмпирической формулой из работ [15, 16] (определяющей остаточную ёмкость в зависимости от времени хранения) при:

$$q(t) = \Delta q_0 e^{-\gamma t} + q_{\Pi} \quad (21)$$

$$\gamma = \frac{1}{Cr}. \quad (22)$$

В данных работах исследовался саморазряд герметичных никель-кадмиевых аккумуляторов.

Формулы (19, 20), согласно (18), справедливы при очень малых u , и, следовательно, из (19) при очень больших сроках хранения. В этом случае формулы (6), (9) вообще неверны. Так как при $t \rightarrow \infty$, согласно (6), логарифм теоретически может возрасти до бесконечности, при этом напряжение на клеммах аккумулятора в результате саморазряда может становиться отрицательным и стремиться к минус бесконечности, что лишено физического смысла. Следовательно, формулы (19), (20) дополняют формулы (6), (9) в указанной области изменения времени хранения.

Таким образом, предложенная структурная модель щелочного аккумулятора с учётом саморазряда позволяет получить все наиболее известные эмпирические зависимости, описывающие процесс саморазряда в щелочных аккумуляторах. Кроме того, из проведённого анализа видно, что различные эмпирические зависимости, описывающие процесс саморазряда, не противоречат, а дополняют друг друга, так как они справедливы каждое в своём интервале рассмотрения времени саморазряда.

Сравнение с экспериментом

Сравним полученные формулы с результатами экспериментальных исследований по саморазряду герметичных никель-кадмиевых аккумуляторов ЦНК-0,45. Результаты экспериментальных исследований по изменению напряжения на клеммах аккумулятора и относительной остаточной ёмкости в процессе саморазряда представлены в таблице 1. Данные результаты во многом схожи с результатами аналогичных исследований из монографии [7].

Таблица 1

Изменение напряжения на клеммах и относительной остаточной ёмкости в процессе саморазряда аккумулятора ЦНК-0,45

Время, сут.	1	3	6	15	30	60
Относительная остаточная ёмкость	0,9	0,879	0,85	0,832	0,809	0,798
Напряжение на клеммах, В	1,314	1,313	1,311	1,31	1,309	1,308

Из таблицы видно, что аккумулятор в первые сутки саморазряда теряет 10 % ёмкости, к концу 30 суток остаточная ёмкость составляет около 80 % и в дальнейшем мало меняется.

Для сравнения полученных формул (6), (9) с экспериментальными данными перепишем их в виде:

$$u_k = E_0 - B1 \cdot \ln[D \cdot t + 1], \quad (23)$$

$$q_{ot} = \frac{B1}{\Psi_0} \ln[D \cdot t + 1], \quad (24)$$

где $B1 = \frac{1}{a}$; $D = \frac{ai_0 \exp(au_0)}{C}$; $q_{ot} = \frac{q}{Q}$; $C = \frac{Q}{\Psi_0}$, Q – полная ёмкость аккумуля-

тора в начале саморазряда; q_{ot} – относительная потеря ёмкости при саморазряде; Ψ_0 – изменение напряжения на линейном участке разрядной кривой аккумулятора. Согласно [7] для аккумулятора ЦНК-0,45 оно равно $\Psi_0 = 0,06$ В.

Найдём коэффициенты E_0 , $B1$, D , A , B соотношений (23) и (8) из требования наилучшего совпадения данных зависимостей с экспериментальными данными из таблицы 1. Для этого будем использовать метод наименьших квадратов и процедуру оптимизации Левенберга-Маркардта. Результаты приведены в таблице 2.

Таблица 2

Оптимальные коэффициенты соотношений (23) и (8)

E_0	$B1$	D	A	B
1,32	$1,611 \cdot 10^{-3}$	37,33	1,314	$1,611 \cdot 10^{-3}$

Относительная ошибка аппроксимации для обеих кривых – менее одного процента. Таким образом, соотношение Гинделиса (8) и соотношение (23) в равной степени хорошо аппроксимируют экспериментальные точки. Первоначально соотношение Гинделиса (8) было получено из анализа саморазряда окисно-никелевого электрода [11], однако, как видно, оно хорошо описывает и саморазряд никель-кадмиевого аккумулятора в целом.

На рисунке 2 приведены графики соотношений (8), (23) с найденными коэффициентами и экспериментальные точки. Как видно из рисунка, графики практически сливаются.

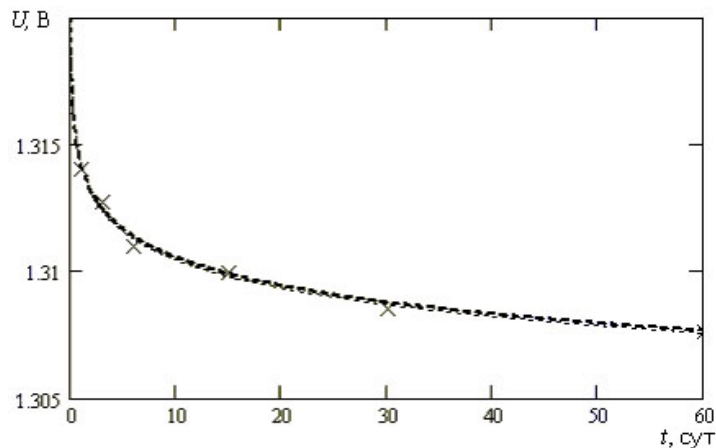


Рис. 2. Изменение напряжения на клеммах аккумулятора ЦНК-0,45 в процессе его саморазряда:

× × – экспериментальные точки, — — кривые, рассчитанные с помощью соотношений (23) и (8) соответственно

Как было показано ранее, соотношения (8) и (23) будут существенно отличаться только при нарушении условия (7). Согласно данным таблицы 2, это будет при

$$t \leq \frac{1}{D} = 0,027 \text{ сут},$$

т.е. в первый час после начала саморазряда, что с практической точки зрения несущественно.

Используя уже найденные коэффициенты из таблицы 2 и соотношение (24), можно произвести независимое сравнение предлагаемой модели саморазряда аккумулятора на базе новых экспериментальных точек по изменению относительной остаточной ёмкости из таблицы 1. Результаты сравнения представлены на рисунке 3. Относительная ошибка аппроксимации также меньше одного процента. Таким образом, предлагаемая модель очень хорошо соответствует экспериментальным данным по саморазряду аккумулятора ЦНК-0,45.

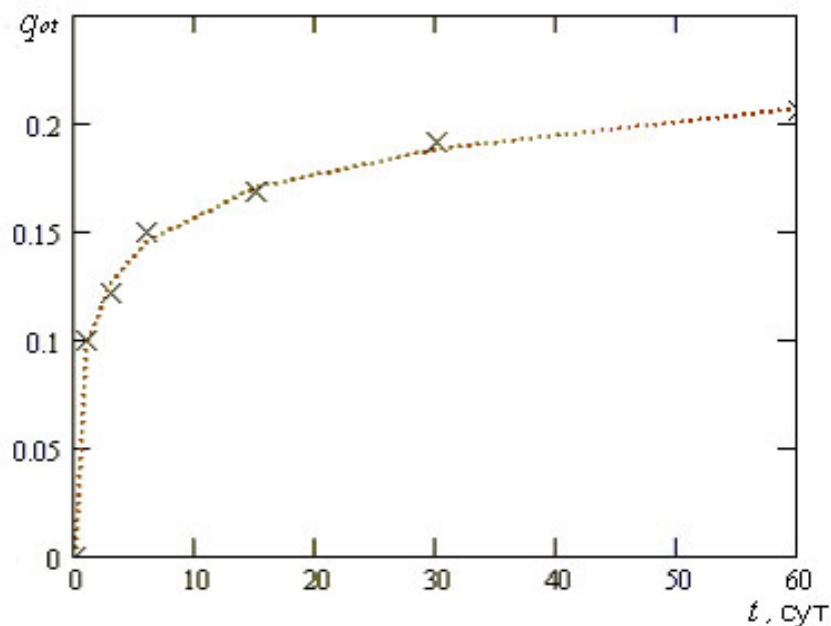


Рис. 3. Изменение относительной потери ёмкости при саморазряде аккумулятора ЦНК-0,45: крестиками отмечены экспериментальные точки, а точками – кривая, рассчитанная с помощью соотношения (24)

В заключение сравним экспериментальные точки с эмпирическим уравнением (21). Используя процедуру оптимизации, описанную выше, можно найти оптимальные коэффициенты γ , $q_{п}$, Δq_0 , однако, если использовать все экспериментальные точки, ошибка аппроксимации будет достаточно большой – около 10 % – и эмпирическая кривая (21) будет плохо соответствовать экспериментальным точкам. Если же не учитывать первые две экспериментальные точки, то относительная ошибка аппроксимации будет также меньше 1 %, при этом оптимальные коэффициенты будут:

$$\gamma=0,049, \quad q_{п}=0,79, \quad \Delta q_0=0,081.$$

Результаты сравнения представлены на рисунке 4. Из графика видно, что эмпирическое уравнение саморазряда (21) хорошо соответствует экспериментальным данным по саморазряду аккумулятора ЦНК-0,45 при времени саморазряда больше 6 суток в соответствии с требованием соотношения (18).

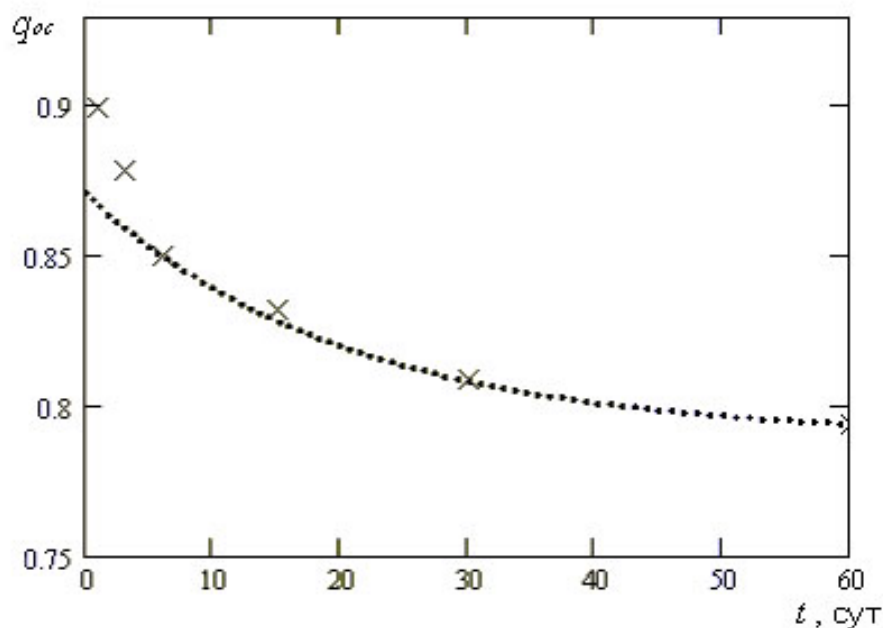


Рис. 4. Изменение относительной остаточной ёмкости при саморазряде аккумулятора ЦНК-0,45: крестиками отмечены экспериментальные точки, а точками – кривая, рассчитанная с помощью соотношения (21)

Так как соотношение (21) справедливо при любых больших временах саморазряда, то найденное $q_{п} = 0,79$ надо считать предельно возможной остаточной ёмкостью для аккумулятора ЦНК-0,45. В этом случае, используя соотношение (24), можно найти ограничение сверху на использование соотношений (23), (24), (8). Действительно, относительная потеря ёмкости при саморазряде аккумулятора ЦНК-0,45 не должна превысить $q = (1 - q_{п}) = 0,21$. Это будет согласно (24) при времени саморазряда 72,8 суток. Таким образом, для аккумулятора ЦНК-0,45 эмпирическое уравнение (8) справедливо в интервале $0,027 < t < 72,8$ суток, т.е. охватывает весь практически интересный интервал саморазряда, а эмпирическое уравнение (21) – в интервале $6 < t < \infty$, т.е. при больших временах саморазряда.

Список литературы

1. Стойнов З.Б. Электрохимический импеданс / Стойнов З.Б., Графов Б.М., Савова-Стойнова Б. [и др.]. – М.: Наука, 1991. – 336 с.
2. Маделунг Э. Математический аппарат физики. – М.: Мир, 1961. – 620 с.
3. Галушкин Н.Е., Галушкина Н.Н. // Электрохимическая энергетика. 2005. – Т. 5, № 1. – С. 43–50.
4. Луковцев П.Д., Темерин С.А. О природе потенциалов и электрохимическое поведение реальных окисных электродов // Совещание по электрохимии; сб. тр. – М.: Изд-во академии наук СССР, 1953. – С. 494–503.

5. Мамыкина Д.А., Абахаев М.Г., Деева Н.С. Об окислении графита в анодной массе работающего щелочного аккумулятора // Исследование в области химических источников тока: сб. работ. – Саратов: Изд-во Саратовского ун-та, 1970. – С. 87–93.
6. Маландин О.Г., Битюцкий П.Н., Тихонова Т.С. Исследование механизма саморазряда металлокерамического окисноникелевого электрода в концентрированных растворах калиевой щелочи // Пятое Всесоюзное совещание по электрохимии: сб. тез. докл. – М.: Изд-во Академии наук СССР, 1974. – С. 465–467.
7. Романов В.В., Хашев Ю.М. Химические источники тока. – М.: Советское радио, 1978. – С. 55.
8. Дергилев А.В., Назарова Т.М. // Ж. прикл. химии. – 1993. – Т. 65, № 8. – С. 1775–1779.
9. Цыганков М.С. Влияние аниона NO_3^- на саморазряд щелочного аккумулятора // Технология производства аккумуляторов: сб. материалов науч.-техн. совещания. – М.: Центральный институт науч.-техн. информации электротехн. промышленности и приборостроения, 1960. – С. 83–85.
10. Чизмаджев Ю.А., Маркин В.С., Чирков Ю.Г. Макрокинетика процессов в пористых средах. – М.: Наука, 1971. – 680 с.
11. Гинделис Я.Е. Саморазряд щелочных аккумуляторов: автореф. дис. ... канд. техн. наук. – Л., 1954. – 25 с.
12. Абрамовиц М., Стиган И. Справочник по специальным функциям. – М.: Наука, 1979. – 830 с.
13. Антропов Л.И. Теоретическая электрохимия. – М.: Высш. шк., 1975. – С. 320.
14. Гинделис Я.Е., Максимова А.А. Температура и кинетика саморазряда окисноникелевого электрода // Исследование в области химических источников тока: сб. работ. – Саратов: Изд-во Саратовского ун-та, 1974. – Вып. 3. – С. 64–67.
15. Клещук В.К., Леонова М.В., Солдатенко Л.А. Динамика газообразных процессов в герметичных НК аккумуляторах: сб. науч. тр. по ХИТ ВНИАМ. – СПб.: Энергоатомиздат, 1991. – С. 45–48.
16. Даниленко И.Ф., Хаскина С.М., Смирнова Т.Н. [и др.]. Математическая модель процесса саморазряда никель-кадмиевых герметичных аккумуляторов в начальный период хранения: сб. науч. тр. по ХИТ. – Л.: Энергоатомиздат, 1989. – С. 87–91.

УДК 004.942

Д.В. Медведев

АЛГОРИТМ ОПРЕДЕЛЕНИЯ ПРОПУСКНОЙ СПОСОБНОСТИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ КОНТРОЛЯ И УЧЁТА ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ В СЕТЯХ 0,4 кВ

В статье рассматривается вопрос оптимизации пропускной способности радиоканала управления ROCSAG для автоматизированной системы контроля и учёта электропотребления (АСКУЭ) в сетях 0,4 кВ. Рассмотрение программно-аппаратного комплекса АСКУЭ как нелинейной динамической системы позволило предложить оптимальный порядок обслуживания, минимизирующий средние стоимостные потери в единицу времени.

Ключевые слова: автоматизированная система контроля и учёта электроэнергии, функция Лагранжа, оператор Урысона, нелинейные динамические системы.

Введение. В современной российской электроэнергетике довольно остро стоит вопрос об эффективном управлении процессом потребления электроэнергии в сетях 0,4 кВ. Процент коммерческих потерь составляет внушительные цифры. Так, например, по данным Шахтинских районных электросетей, в 2009 г. он стабильно превышал 30 % от общего объёма отпущенной потребителям электроэнергии.

Существующее положение усугубляется отсутствием возможности оперативного адресного воздействия на неплательщиков. Также представляется затруднительным выявление хищений электроэнергии путём периодических обходов абонентов контролёрами электросетей, что подтверждают приведённые выше данные.

Решением описанной проблемы является внедрение предприятиями энергосбыта автоматизированных систем контроля и учёта электроэнергии (АСКУЭ). Анализ существующих АСКУЭ позволяет сделать вывод, что оптимальной, с точки зрения экономической эффективности и надёжности, является АСКУЭ с дистанционным управлением по радиоканалу.

Постановка задачи

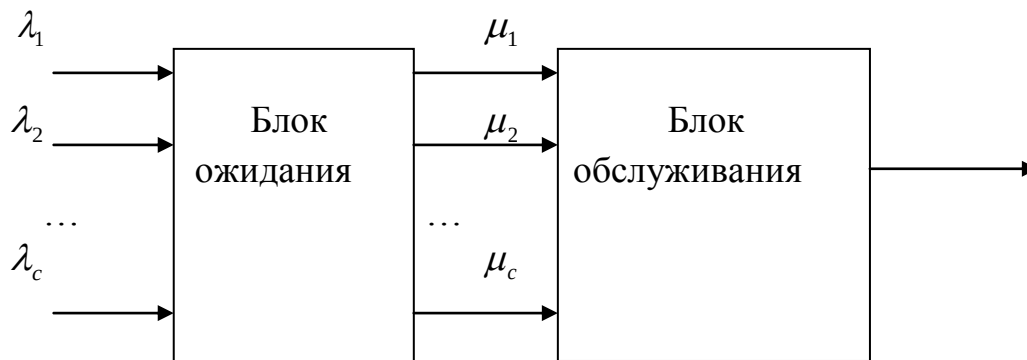
Рассмотрим ряд вопросов, связанных с проектированием АСКУЭ, среди которых одним из главных является вопрос о пропускной способности такой системы.

Как известно, стоимостные модели теории массового обслуживания ориентированы на определение оптимального уровня обслуживания, при котором должен достигаться необходимый баланс между прибылью, получаемой за счёт работы системы массового обслуживания, и потерями, а также затратами фирмы, связанными с задержками в предоставлении услуг и в нормальном обеспечении ресурсами.

В самом простом случае рассматривают модель типа $(M/M/1):(GD/\infty/\infty)$ с одним обслуживающим узлом, со средней частотой (скоростью) поступления заявок, равной $\lambda > 0$, со средней скоростью обслуживания узла, равной $\mu > 0$.

В рассматриваемой нами АСКУЭ действует несколько узлов (операторов) и несколько каналов. С этой целью рассмотрим многоканальную модель типа $(M/M/c):(GD/\infty/\infty)$, где c – число обслуживающих узлов и соответствующее число каналов больше единицы.

Введём необходимые для дальнейшего изложения обозначения. Обозначим через C_{1k} выигрыш в денежных единицах за счёт увеличения на одну единицу значения μ_k в течение единичного интервала времени, где k – номер узла, $k = \overline{1, c}$.



:Пусть C_{2k} – экономические потери в единицу времени на k -м узле из-за вынужденного ожидания одного клиента (заявки).

Средняя продолжительность обслуживания на k -м узле – $\frac{1}{\mu_k}$, средняя скорость поступления заявок на k -м узле равна λ_k , среднее число клиентов, находящихся в k -м канале системы, обозначим через L_{sk} , где $k = \overline{1, c}$. Считаем, что одновременно в рассматриваемой системе массового обслуживания действуют независимые c операторов, следовательно

$$L_{sk} = \frac{\rho_k}{1 - \rho_k},$$

где $\rho_k = \lambda_k / \mu_k$.

Таким образом,

$$L_{sk} = \frac{\lambda_k}{\mu_k - \lambda_k}, \quad k = \overline{1, c}.$$

Ограничения по скорости обслуживания в целом по системе будут равны:

$$\sum_{k=1}^c a_k \mu_k \leq A,$$

где a_k – весовой коэффициент ограничения k -го узла.

Аналогично, как в одноканальной системе массового обслуживания, можно вывести выражение целевой функции многоканальной модели, которое принимает следующий вид:

$$F(\mu_1, \mu_2, \dots, \mu_c) = \sum_{k=1}^c (C_{1k} \cdot \mu_k + C_{2k} \cdot L_{sk}),$$

где $L_{sk} = \lambda_k / (\mu_k - \lambda_k)$, для $k = \overline{1, c}$.

Данная условно экстремальная задача имеет следующие ограничения:

$$\sum_{k=1}^c a_k \mu_k \leq A.$$

Пути решения задачи

Решение поставленной условно экстремальной задачи можно найти, используя так называемую функцию Лагранжа и метод неопределённых множителей Лагранжа.

Сначала строится функция Лагранжа, имеющая вид:

$$L(\mu_1, \mu_2, \dots, \mu_c) = F(\mu_1, \mu_2, \dots, \mu_c) + \alpha \cdot \Phi(\mu_1, \mu_2, \dots, \mu_c), \quad (1)$$

где F – целевая функция модели; Φ – функция ограничений; α – так называемый неопределённый множитель Лагранжа.

Оптимальные значения для скорости обслуживания на каждом узле, а также оптимальные значения множителя α найдём, продифференцировав по $\mu_1, \mu_2, \dots, \mu_c$ функцию Лагранжа (1) и приравнявая их к нулю.

Получим следующую систему алгебраических уравнений:

$$\begin{cases} L'_{\mu_k} = C_{1k} - C_{2k} \cdot \lambda_k / (\mu_k - \lambda_k)^2 - \alpha \cdot a_k = 0 \\ L'_{\alpha} = A - \sum_{k=1}^c a_k \cdot \mu_k \end{cases},$$

где $k = \overline{1, c}$

Отсюда получаем:

$$C_{1k} - \alpha \cdot a_k = \frac{C_{2k} \lambda_k}{(\mu_k - \lambda_k)^2};$$

$$\mu_k - \lambda_k = \sqrt{\frac{C_{2k} \lambda_k}{C_{1k} - \alpha \cdot a_k}}.$$

Аналогично одноканальной системе массового обслуживания получим расчётные формулы для скоростей обслуживания на каждом из узлов системы:

$$\mu_k^* = \lambda_k + \sqrt{\frac{C_{2k} \lambda_k}{C_{1k} - \alpha \cdot a_k}}, \quad (2)$$

где $k = \overline{1, c}$.

Как видим, данные расчётные формулы аналогичны полученной ранее, если считать, что коэффициент (множитель Лагранжа) $\alpha = 0$.

Однако в общем случае формулы (2) зависят от параметра α .

Чтобы определить требуемое значение α^* , подставим найденные оптимальные значения $\mu_1^*, \mu_2^*, \dots, \mu_c^*$ в ограничения задачи:

$$\sum_{k=1}^c a_k \cdot \mu_k^* = A. \quad (3)$$

Здесь все величины, кроме α , нам известны. Решив уравнение (3) относительно неизвестного α , получим требуемое оптимальное решение поставленной задачи.

Остановимся вкратце на некоторых чаще встречающихся на практике нелинейных динамических системах. В случае модели нелинейной системы её структуру, как правило, представляют в виде комбинаций более простых звеньев: линейного и функционального, предполагая, что в функциональном звене длительность переходного процесса пренебрежимо мала. Нелинейные операторы, входящие в модель, – это, как правило:

- нелинейный оператор суперпозиции Немыцкого $y(\tau) \equiv \Phi(\tau, x(\tau))$;
- нелинейный интегральный оператор Гаммерштейна:

$$H(x(\tau)) \equiv \int_0^T K(\tau, t_1) \Phi(t_1 x(t_1)) dt_1;$$

- оператор Урысона: $I(x(\tau)) \equiv \int_0^T K(\tau, \tau_1) x(\tau_1) d\tau_1$;

- оператор Лихтенштейна – Ляпунова:

$$L(x(\tau)) \equiv \sum_{p_1, p_2, \dots, p_k=0}^N \int_0^T \dots \int_0^T K_{p_1, p_2, \dots, p_k}(\tau, \tau_1, \tau_2, \dots, \tau_k) \cdot x^{p_1}(\tau_1) x^{p_2}(\tau_2) \dots x^{p_k}(\tau_k) d\tau_1 d\tau_2 \dots d\tau_k.$$

Естественно, что не все нелинейные системы поддаются описанию моделями с помощью операторов, связывающих в явном виде выход системы с её входом. Если, например, в рассматриваемой нелинейной системе электроснабжения существует обратная связь, то эта система описывается операторным уравнением вида:

$$y(\tau) = B(x(\tau), y(\tau)) + x(\tau),$$

где B – нелинейный оператор.

В системах контроля и учёта электроэнергии, являющихся следящими, оператор B часто удаётся представить в виде нелинейного интегрального Гаммерштейна H .

Нелинейные динамические системы, описываемые нелинейными обыкновенными дифференциальными уравнениями, целесообразно сводить к отвечающим им нелинейным интегральным уравнениям [3]. Возможные приложения динамических характеристик (т.е. операторов, устанавливающих зависимость выхода от входа в систему в любой момент времени): задачи синтеза систем; задачи управления при оценке основных

параметров; для более глубокого понимания сути технологического процесса, особенно при построении динамических характеристик системы аналитическими методами; для получения статистических характеристик; для нахождения спектральных свойств выходного сигнала по известным спектральным свойствам входного сигнала, и наоборот; для понимания характера протекания переходных процессов и оценок времени запаздывания, а также времени окончания переходных процессов по различным каналам нелинейной системы обслуживания с приоритетами. В последнем из рассматриваемых приложений в систему обслуживания поступает несколько независимых в совокупности пуассоновских потоков потребления электроэнергии, где заявки k -го потока последовательно проходят обслуживание по k -й цепочке. У каждого из датчиков допускается ожидание; в каждый момент обслуживание происходит не более чем на одном канале; прерывание обслуживания не допускается.

Оптимальный порядок обслуживания (минимизирующий средние стоимостные потери в единицу времени в стационарном режиме) можно вычислить, используя методику, изложенную в работе [4]. Аналогичные соображения для расчёта оптимального порядка обслуживания и последующего контроля потреблённой электроэнергии для систем с прерыванием (отказами) приведены в монографии [5].

Заключение

Рассмотрение активной АСКУЭ с функцией управления по радиоканалу [4] как нелинейной динамической системы позволяет определить оптимальный порядок обслуживания, минимизирующий средние стоимостные потери в единицу времени. Такой подход, несомненно, будет полезен при проектировании и внедрении подобных систем на практике.

Список литературы

1. Рыжиков Ю.И. Теория очередей и управление запасами. – СПб.: Питер, 2001. – 384 с.
2. Фетисов В.Г., Сапронов А.А., Медведев Д.В. Аналитические методы в нелинейных динамических системах контроля и учёта электроэнергии // Математические методы в технике и технологиях: сб. тр. XVI Междунар. науч. конф.: в 10 т. Т. 8. Секция 12 / под общ. ред. В.С. Балакирева. – Ростов н/Д.: РГАСМ ГОУ, 2003. – С. 52–54.
3. Фетисов В.Г. Операторы и уравнения в F -квазинормированных пространствах : дис. ... докт. физ.-мат. наук / Ин-т матем. им. акад. Соболева С.Л. Сиб.отд. РАН. – 1996. – 280 с.
4. Медведев Д.В. Математическое моделирование и алгоритмы функционирования автоматизированной системы контроля и учёта электроэнергии в сетях О, U и В : дис. ... канд. техн. наук / Южно-Рос. гос. техн. ун-т (НПИ). – 2005. – 139 с.
5. Иванов В.В. Методы алгоритмизации непрерывных производственных процессов. – М.: Физматгиз, 1975. – 400 с.

УДК 621.01:004.4

Ю.А. Валюкевич, А.А. Зеленский

ПРОГРАММНЫЙ ИНТЕРПОЛЯТОР ОТРЕЗКА ПРЯМОЙ ДЛЯ СИСТЕМЫ УПРАВЛЕНИЯ ДВУХЗВЕННЫМ МЕХАНИЗМОМ

В работе приведены результаты проектирования алгоритма интерполяции отрезка прямой, произвольно расположенного на плоскости прямой, путём преобразования двух вращательных движений звеньев двухподвижного механизма.

При обработке плоских заготовок практически не используются механические системы, формирующие траекторию перемещения обрабатывающего инструмента как сумму двух вращательных движений. В самом деле, такой способ формирования произвольной программно заданной траектории имеет ряд существенных недостатков: переменная величина минимальной ошибки позиционирования рабочего инструмента в заданную точку; достаточно высокая сложность поддержания постоянства контурной скорости (весьма важное свойство механизма для большинства практических приложений); отсутствие обобщённого подхода к решению задачи интерполяции кривой для неортогональных механизмов; высокая сложность программно-аппаратной системы управления движением по заданному контуру рабочего инструмента.

Однако механизм преобразования двух вращательных движений в программно-заданную траекторию позволяет существенно упростить конструкцию технологического оборудования, снизить его материало- и энергоёмкость. При этом предполагается существенное усложнение интерполяционных алгоритмов и системы управления скоростями вращения звеньев механизма [1].

Кинематическая схема содержит круглый стол, вращающийся относительно вертикальной оси, и инструментальную штангу, вращающуюся вокруг параллельной оси на расстоянии p с закреплённым на конце штанги в точке M инструментом (рис. 1).

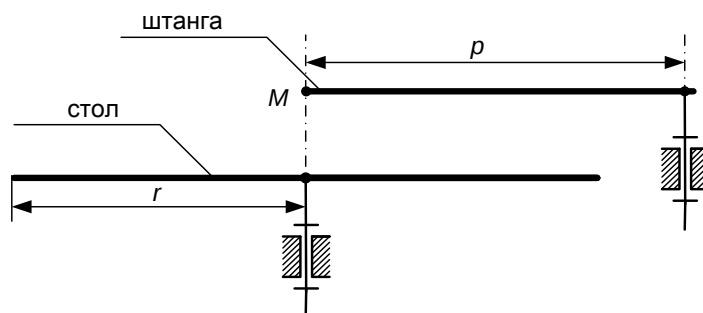


Рис. 1. Кинематическая схема устройства (вид сбоку)

На рисунке 1 приняты следующие обозначения: r – радиус рабочего стола; p – длина инструментальной штанги.

При использовании в системе управления движением звеньев двухкоординатного механизма вычислителя с достаточно высокой производительностью интерполяционный алгоритм можно реализовать на основе совместного решения уравнения прямой, которой принадлежит интерполяционный отрезок, и уравнения окружности, образованной на поверхности рабочего стола концом инструментальной штанги (рабочего инструмента). Для реализации методики решения интерполяционной задачи определим начальные условия:

- контур обрабатываемой детали задан в виде последовательности отрезков прямых линий и описан массивом полярных координат начал и концов отрезков $\{\rho_j, \alpha_j\}$, $j = 0, 1, 2, \dots, m$;

- все расчёты ведутся в неподвижной декартовой системе координат. Причём начало декартовой и полюс полярной системы координат находятся на оси вращения стола, а оси абсцисс и нулевого радиус-вектора совпадают и проходят через центр стола и ось вращения штанги;

- ведущей координатой при интерполяции при любом положении отрезка на поверхности рабочего стола является угловая координата стола α ;

- значения шага интерполяции по обеим угловым координатам $\Delta\alpha$ и $\Delta\beta$ равны между собой.

На рисунке 2 приведена геометрическая иллюстрация, поясняющая выполнение одного шага интерполяционного алгоритма.

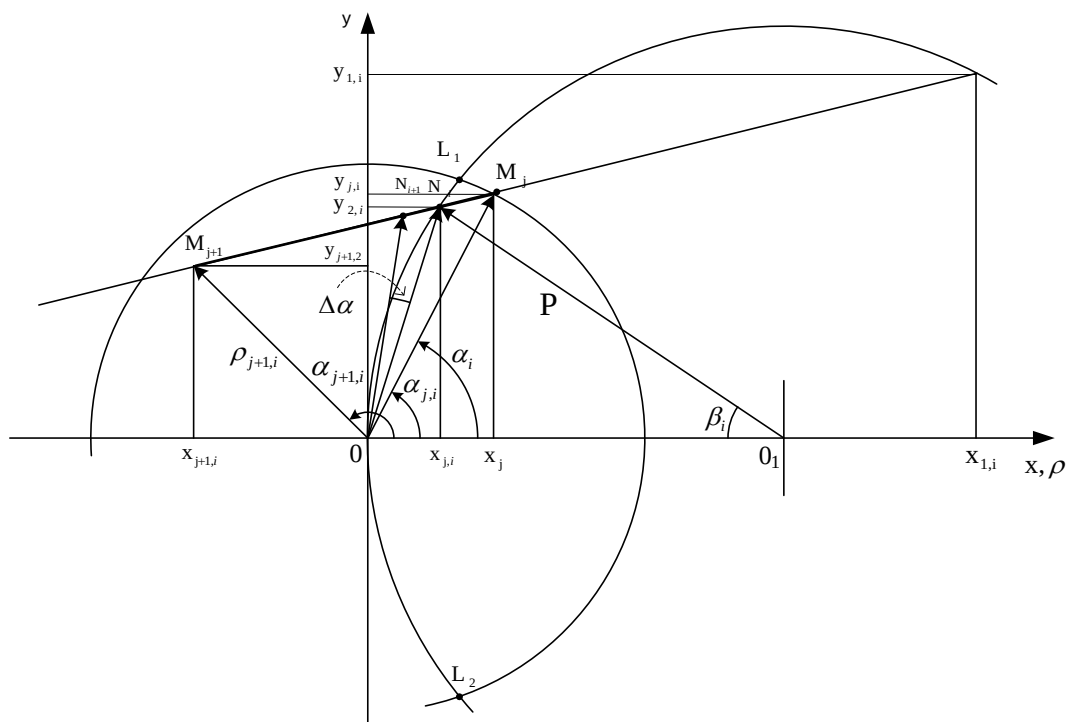


Рис. 2. Геометрическая интерпретация интерполяционного алгоритма методом решения системы уравнений

На рисунке 2 приняты следующие обозначения:

i – индексная переменная номера интерполяционной точки $i = 0, 1, 2 \dots n$;

j – индексная переменная номера интерполяционного отрезка $j = 0, 1, 2 \dots m$;

$x_{j,i}, y_{j,i}$ – координаты начала и конца отрезка в начальной точке интерполяционного участка;

$\rho_{j,i}, \alpha_{j,i}, \rho_{j+1,i}, \alpha_{j+1,i}$ – полярные координаты начала и конца отрезка, соответствующие декартовым, где α_i, ρ_i – угловые координаты начала интерполируемого участка;

$x_{1,i}, x_{2,i}, y_{1,i}, y_{2,i}$ – корни системы уравнений прямой и дуги окружности, причём $x_{2,i}, y_{2,i}$ соответствуют координатам начальной точки интерполируемого участка;

$\Delta\alpha$ – шаг интерполяции по угловой координате стола;

β_i – текущее угловое положение штанги;

R – радиус штанги;

N_i, N_{i+1} – точки начала и конца отрезка, соответствующего изменению угловой координаты стола на $\Delta\alpha$ (угловой шаг интерполяции).

Необходимо отметить, что показанное положение точек N_i, N_{i+1} не учитывает линейную ошибку интерполяции и является идеализированным представлением процесса.

Исходное описание контура может быть представлено в виде текстового файла в формате *.plt или в формате так называемых G-кодов, применяемых в станках с ЧПУ. Преобразование подобного представления контура в массив вида $\{\rho_j, \alpha_j\}$ существенных трудностей не представляет (рис. 3).

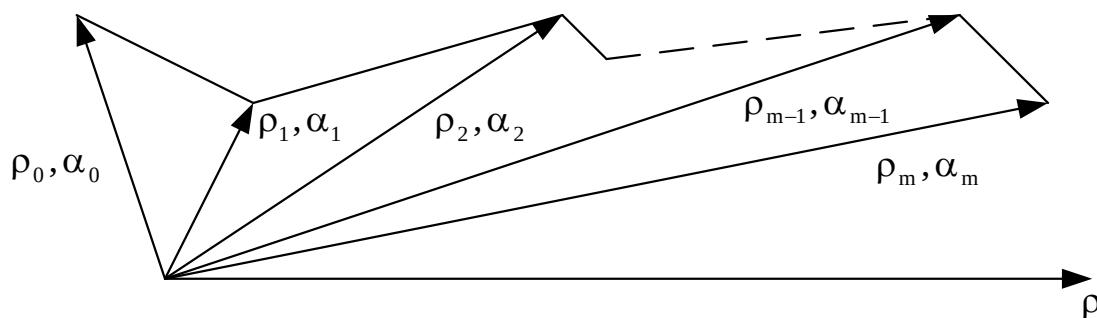


Рис. 3. Контур обработки детали

Блок-схема алгоритма интерполяции отрезка прямой приведена на рисунке 4. Исполнение алгоритма начинается с ввода элементов массива описания контура очередной пары элементов ρ_{i+1}, α_{i+1} , описывающих исходное (до начала интерполяции контура) положение конца отрезка (блок 1 алгоритма).

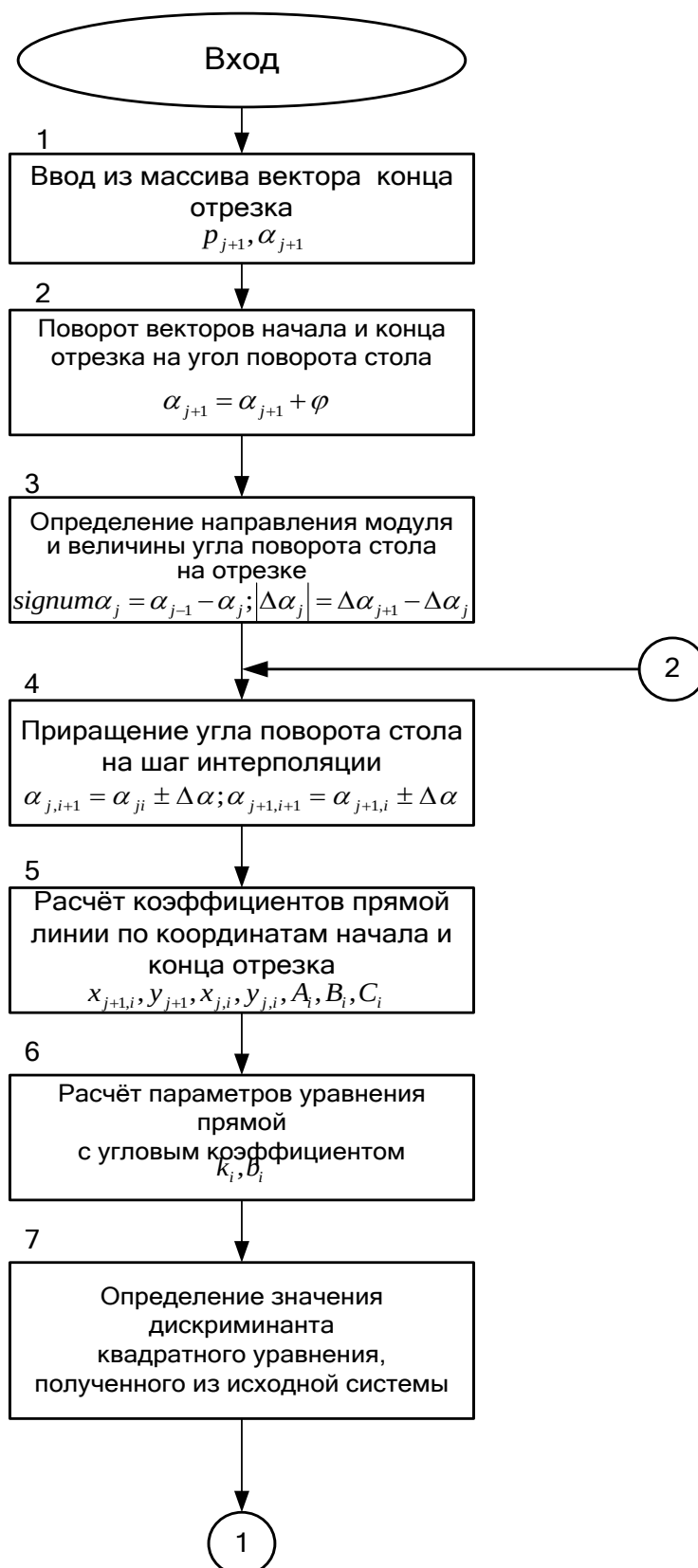


Рис. 4. Схема алгоритма интерполяции отрезка прямой методом решения системы уравнений (начало; продолжение и окончание см. на с. 158 и 159)

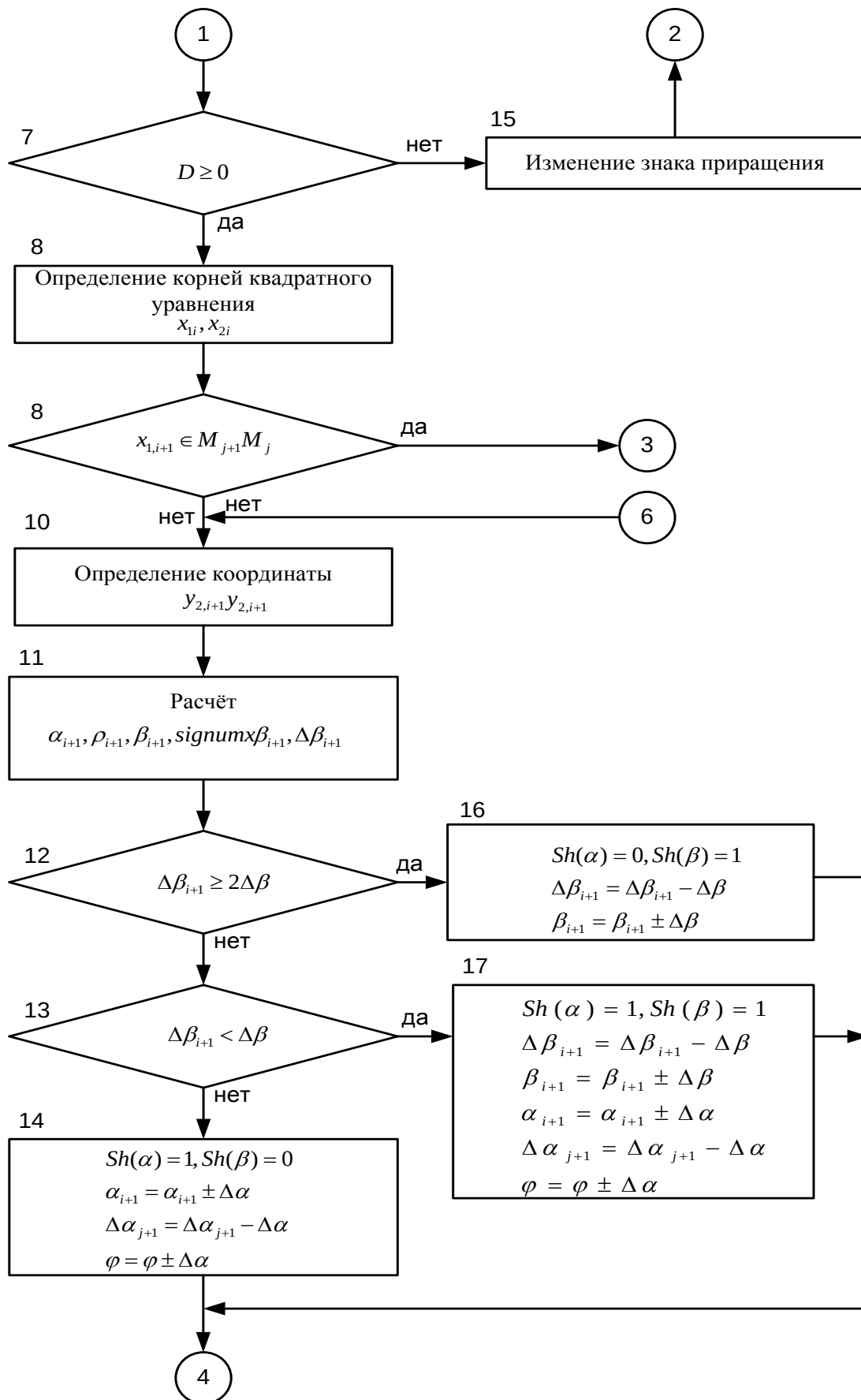


Рис. 4. Продолжение (начало см. на с. 157, окончание – на с. 159)

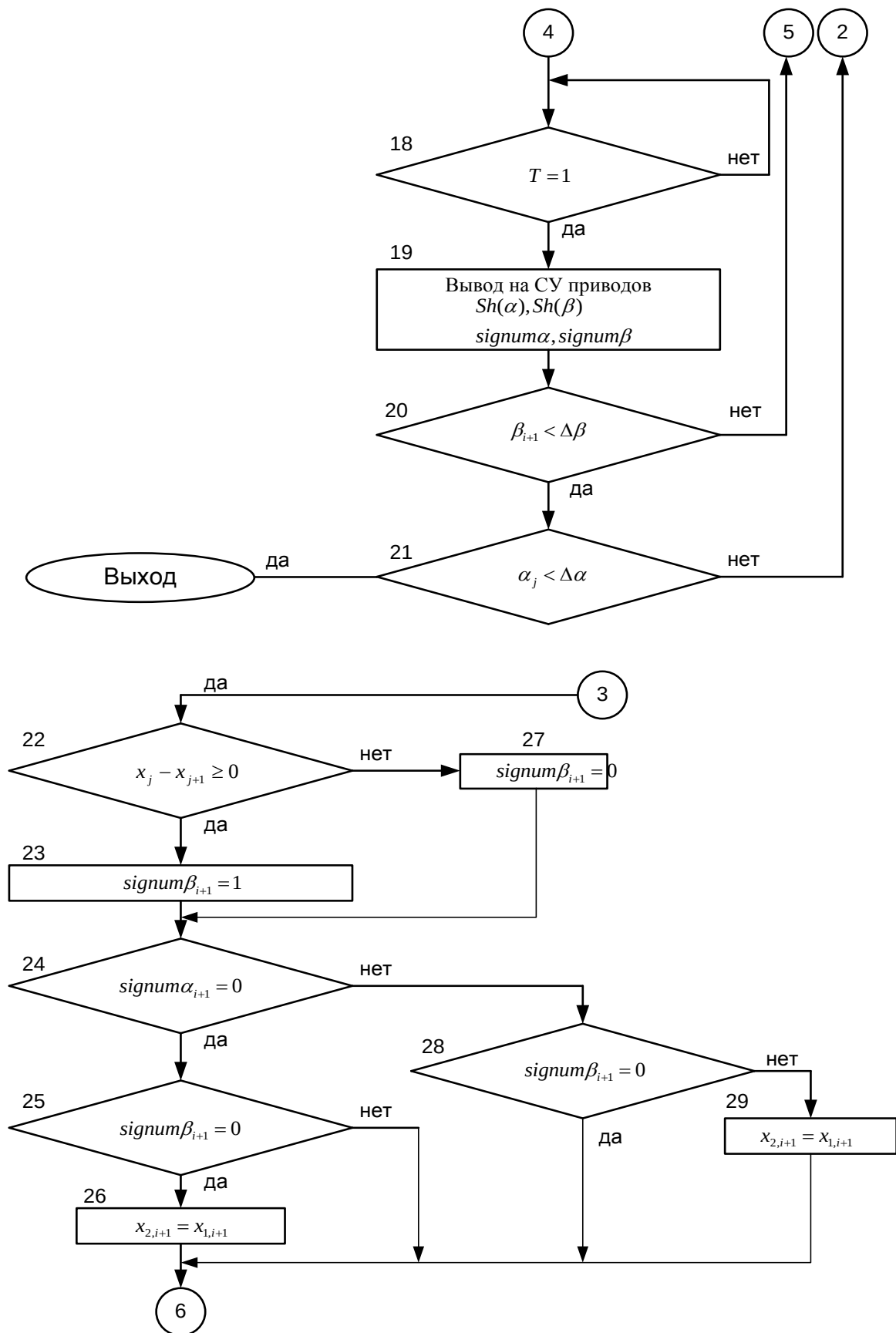


Рис. 4. Окончание (начало см. на с. 150 и с. 151)

Для получения полярной координаты действительного положения текущего конца отрезка необходимо учесть угол поворота стола при интерполяции предыдущих отрезков прямой. Предполагается, что на всём протяжении работы устройства текущий угол поворота стола φ определяется независимо от режима работы. При этом полярная координата конца текущего отрезка при известном угловом положении стола определяется как $\alpha_{j+1} = \alpha_{j+1} + \varphi$ (блок 2 алгоритма).

Координата α_j начала отрезка известна из предыдущего этапа интерполяции, т.е. является одновременно концом предыдущего отрезка и находится в исходной точке рабочей поверхности стола для текущего этапа. Направление изменения угловой координаты на текущем участке определяется как $\text{sign } \alpha_j = \alpha_{j+1} - \alpha_j$ (блок 3 алгоритма). Здесь же определяется модуль угла $\Delta\alpha_j$, соответствующий угловому приращению текущего отрезка $\Delta\alpha_j = \alpha_{j+1} - \alpha_j$ (блок 3 алгоритма).

Дальнейшее описание работы алгоритма интерполяции приведём на примере формирования заданий для привода по угловым координатам α и β на примере двух соседних точек, принадлежащих отрезку M_1M_2 и расположенных на некотором удалении от начала отрезка (рис. 1). Первым шагом собственно интерполяционного алгоритма является изменение ведущей координаты α в зависимости от $\text{sign } \alpha_j$ на угол $\pm\Delta\alpha$. Интерполяционный алгоритм построен на решении системы двух уравнений, первое из которых – уравнение отрезка прямой в системе координат, описанной выше, после приращения угла поворота стола, второе – уравнение дуги окружности, описываемой концом инструментальной штанги (L_1L_2 , рис. 1).

Эта система может быть представлена в полярной системе координат в виде:

$$\begin{cases} \rho(A_{i+1} \cos \alpha_{i+1} + B_{i+1} \sin \alpha_{i+1}) + C_{i+1} = 0 \\ \rho^2 - 2\rho P \cos \alpha_{i+1} = 0 \end{cases} \quad (1)$$

или в декартовой системе координат:

$$\begin{cases} y = k_i x + b_c \\ (x - P)^2 + y^2 = P^2. \end{cases} \quad (2)$$

Решение системы (2) является более предпочтительным, т.к. алгоритм численного решения этой системы проще, чем системы (1). Для определения коэффициентов первого уравнения системы (2) необходимо определить истинное положение концов отрезка после поворота на угол $\Delta\alpha$, например, используя параметрические уравнения прямой для извест-

ных $\rho_{j,i}, \alpha_{j,i}, \rho_{j+1,i}, \alpha_{j+1,i}$, а затем при известных координатах начала и конца отрезка, воспользовавшись уравнением прямой линии в приращениях, которой принадлежит отрезок

$$\frac{x - x_{j,i}}{x_{j+1,i} - x_{j,i}} = \frac{y - y_{j,i}}{y_{j+1,i} - y_{j,i}}, \quad (3)$$

можно определить коэффициенты уравнения прямой линии (блок 5 алгоритма):

$$\begin{aligned} A_i &= y_{j+1,i} - y_{j,i}; & B_i &= x_{j+1,i} - x_{j,i}; \\ C &= x_{j,i}(y_{j+1,i} - y_{j,i}) - y_{j,i}(x_{j+1,i} - x_{j,i}) \end{aligned} \quad (4)$$

и коэффициенты k_i, b_i первого уравнения системы (блок 6 алгоритма):

$$k_i = \frac{A_i}{B_i}; \quad b_i = -\frac{C_i}{B_i}.$$

Подставляя из первого уравнения системы (2) выражение для y во второе и проводя элементарные преобразования, получим уравнения вида:

$$x^2 - a_{1,i+1}x + a_{2,i+1} = 0, \quad (5)$$

где

$$a_{1,i+1} = \frac{\rho - k_i b_i}{1 + k_i^2}; \quad a_{2,i+1} = \frac{b_i^2}{1 + k_i^2}.$$

Если дискриминант квадратного уравнения (5) $D = a_{1,i+1}^2 - a_{2,i+1}$ меньше нуля, то это означает, что система (3.13) не имеет решения в области действительных чисел, т.е. точка отрезка N_{i+1} не принадлежит дуге $L_1 L_2$. В этом случае меняется знак приращения, т.е. направление вращения стола, текущее, а значение углов $\alpha_{j,i+1}, \alpha_{j+1,i+1}$ изменяется на предыдущее (блок 15 алгоритма), и осуществляется переход на начало процедуры интерполяции (блок 4 алгоритма).

При $D > 0$ определяются корни квадратного уравнения (5) $x_{1,i+1}$ и $x_{2,i+1}$ и принадлежность этих корней отрезку M_j, M_{j+1} (блок 8 алгоритма). Для большого числа случаев расположения отрезков на поверхности стола первый (большой) корень $x_{1,i+1}$ не принадлежит интерполируемому отрезку и поэтому осуществляется переход к блоку 10 алгоритма, где определяется вторая переменная исходной системы уравнений $y_{2,i+1}$.

Полярные координаты точки N_{i+1} определяются из соотношений (блок 11 алгоритма):

$$\rho_{i+1} = \sqrt{x_{2,i+1}^2 + y_{2,i+1}^2},$$

$$\alpha_{i+1} = \arccos \frac{x_{2,i+1}}{\rho_{i+1}}.$$

Угол поворота штанги, соответствующий значению α_{i+1} , определяется из треугольника $\Delta N_{i+1} x_{2,i+1} 0$ при условии, что точка N_{i+1} принадлежит дуге u, u_2 (на рис. 1 не показано), по формуле:

$$\beta_{i+1} = \pi - \arccos \frac{P - x_{2,i+1}}{P}.$$

Значение $\Delta\beta_{i+1} = \beta_i - \beta_{i+1}$ может существенно превосходить заданный шаг интерполяции $\Delta\beta$, а по условию поддержания заданной контурной скорости приращение за один интерполяционный шаг не может превышать $\Delta\beta$. Необходимо проверить выполнение ряда дополнительных условий, первое из которых – $\Delta\beta_{i+1} > 2 \cdot \Delta\beta$ (блок 12 алгоритма). При этом двоичная переменная $Sh(\alpha)$, соответствующая выдаче управляющего сигнала на привод о текущем единичном перемещении, принимается равной 0 (нет перемещения) $\Delta\beta_{i+1} = \Delta\beta_{i+1} - \Delta\beta$, $\beta_i = \beta_i \pm \Delta\beta$ (блок 15 алгоритма). В случае $\Delta\beta_{i+1} > \Delta\beta$ (блок 13 алгоритма) выполняются действия, представленные в блоке 17 алгоритма, в противном случае – в блоке 14 алгоритма. После расчёта параметров шага интерполяции происходит переход в режим ожидания окончания временного интервала, задающего шаг интерполяции по времени (блок 18 алгоритма). Параметр T определяется с помощью отдельной подпрограммы и задаёт контурную скорость обработки изделия. При $T = 1$ на приводы угловых координат α , β выдаются значения угловых перемещений (есть/нет) и направление перемещения (блок 19 алгоритма). После выдачи управляющих воздействий на приводы осуществляется проверка разности между текущим и заданным положениями штанги (блок 20 алгоритма). Если требуемое перемещение больше шага интерполяции по углу β , то осуществляется переход к блоку 12 алгоритма без изменения угла α . В противном случае осуществляется проверка на достижение конечной точки текущего отрезка (блок 21 алгоритма) и, если условие выполняется, осуществляется выход из программы интерполяции текущего отрезка контура, в противном случае осуществляется переход к блоку 4 алгоритма и процесс интерполяции продолжается.

Если условие в блоке 8 алгоритма истинно, т.е. оба корня уравнения принадлежат интерполируемому отрезку, необходимо определить, какой из корней уравнения должен быть взят для последующих расчётов. Ситуация, возникшая в этом случае, представлена на рисунке 5.

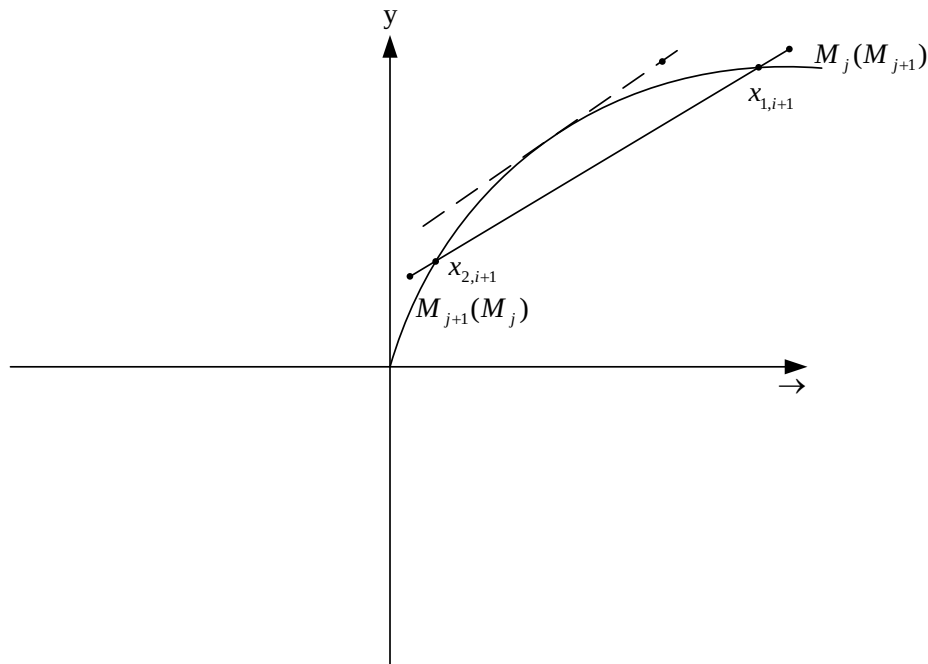


Рис. 5. Геометрическая интерпретация процесса интерполяции
при $x_{1,i+1} \in M_j M_{j+1}$; $x_{2,i+1} \in M_j M_{j+1}$

При ситуации, представленной на рисунке 5, в отличие от рассмотренной выше, необходим дополнительный анализ на соотношение направлений перемещений по угловым координатам, т.е. в расчёт необходимо принимать взаимное положение начальной и конечной точки отрезка и положение корней характеристического уравнения. При условиях $\text{sign } \alpha = 0$ перемещение происходит против часовой стрелки, $\text{sign } \alpha = 1$ – по часовой (то же самое для $\text{sign } \beta$). Здесь необходимо отметить, что направление перемещения по координате β в этом случае не изменяется на всём этапе интерполяции. Выбор корня квадратного уравнения при этом осуществляется в соответствии с блоками 22–28 алгоритмов.

Представленный алгоритм интерполяции отрезка прямой был реализован в виде программы на языке C++, которая работает в режиме моделирования процесса с выводом информации в виде анимационной картинки на экран дисплея и(или) в режиме управления реальным устройством. Разработанная программа может быть достаточно просто портирована на другую программно-аппаратную платформу и работать в режиме реального времени.

Список литературы

1. Зеленский А.А., Валюкевич Ю.А., Кузнецов С.А. Устройство для раскроя плоских материалов с поворотным столом // Известия вузов. Северо-кавказский регион. Машиностроение. – 2008. – № 4. – С. 102–103.

УДК 621.01:004

И.И. Наумов

КИНЕМАТИЧЕСКАЯ МОДЕЛЬ МАНИПУЛЯТОРА ДЛЯ СИЛОМОМЕНТНОЙ ОБРАБОТКИ ПЛОСКИХ МАТЕРИАЛОВ

В работе предложен станок для силомоментной обработки плоских материалов на базе двухзвенного механизма, непосредственно преобразующего два вращательных движения в перемещение инструмента по заданному контуру.

Кинематическая схема станка на базе двухзвенного механизма состоит из двух кругов одинакового радиуса: первый круг вращается относительно вертикальной оси, расположенной в начале координат (рис. 1), а второй закреплён с возможностью вращения вокруг точки крепления H , расположенной на краю первого круга, и рабочей точки $P(t)$, расположенной на внешнем диаметре второго круга и проходящей через центр первого круга.

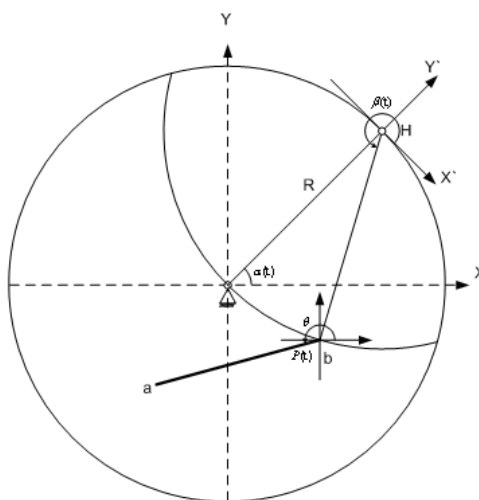


Рис. 1. Кинематическая схема устройства

На рисунке 1 приняты следующие обозначения:

$P(t)$ – рабочая точка с координатами $x(t), y(t)$ в основной (неподвижной) системе координат xu и $x_2(t), y_2(t)$ в дополнительной системе координат $x'y'$;

$\alpha(t)$ – угол поворота диска относительно основной системы координат;

$\beta(t)$ – угол поворота штанги относительно дополнительной системы координат;

a, b – концы отрезка ab , который должна проходить точка $P(t)$;

θ – угол наклона отрезка ab ;

R – радиус дисков.

Поскольку исследуется кинематическая модель двухзвенного манипулятора в двумерной системе, то для определения положения звеньев целесообразнее найти аналитические выражения, связывающие декартовы координаты рабочей точки с изменением полярных координат каждого звена. За основу решения взят метод Денавита – Хартенберга с применением дополнительной системы координат $x'y'$. Так как звеньев в кинематической схеме два, то вместо матриц перехода от одного звена к другому целесообразнее воспользоваться тригонометрической записью [1].

Запишем координаты рабочей точки $P(t)$ относительно осей $x'y'$.

$$\begin{aligned}x_2(t) &= R \cos \beta(t) \\ y_2(t) &= R \sin \beta(t)\end{aligned}\quad (1)$$

Теперь, воспользовавшись формулой перехода из одной системы координат в другую

$$\begin{aligned}x &= x_2 \cos \lambda - y_2 \sin \lambda + a \\ y &= x_2 \sin \lambda + y_2 \cos \lambda + b,\end{aligned}\quad (2)$$

получим

$$\begin{aligned}x &= x' \cos \left(\alpha - \frac{\pi}{2} \right) - y' \sin \left(\alpha - \frac{\pi}{2} \right) + R \cos \alpha \\ y &= x' \sin \left(\alpha - \frac{\pi}{2} \right) + y' \cos \left(\alpha - \frac{\pi}{2} \right) + R \sin \alpha\end{aligned}\quad (3)$$

или

$$\begin{aligned}x(t) &= R \cos \beta(t) \sin \alpha(t) + R \sin \beta(t) \cos \alpha(t) + R \cos \alpha(t) \\ y(t) &= R \sin \beta(t) \sin \alpha(t) - R \cos \beta(t) \cos \alpha(t) + R \sin \alpha(t).\end{aligned}\quad (4)$$

Решая данную систему уравнений, получим:

$$\begin{cases} \alpha(t) = \text{Arc cos} \left(\frac{x(t) \sin \beta(t) + x(t) - \cos \beta(t) y(t)}{2R(\sin \beta(t) + 1)} \right) \\ \beta(t) = \text{Arc sin} \left(\frac{x^2(t) + y^2(t)}{2R^2} - 1 \right) \end{cases}, \quad (5)$$

т.к. $\beta(t)$ может принимать любое значение (ограничения в нашем случае накладываются только областью изменения $x(t)$ и $y(t)$), то любому значению аргумента арксинуса будут соответствовать два угла, вычисляемых по формуле для периода арксинуса:

$$\text{Arc sin} \left(\frac{x^2(t) + y^2(t)}{2R^2} - 1 \right) = (-1)^n \arcsin \left(\frac{x^2(t) + y^2(t)}{2R^2} - 1 \right) + \pi n,$$

где $n=0,1,2,\dots$

Определимся с законами изменения $x(t)$ и $y(t)$. В общем случае их зависимость от времени может носить произвольный характер и не влияет на найденные соотношения. Предположим, что рабочая точка должна двигаться по прямой ab с постоянной линейной скоростью v . Вектор скорости в таком случае совпадает с направлением движения рабочей точки, и пройденное расстояние $S(t) = vt$.

Уравнения $x(t)$ и $y(t)$ можно записать

$$\begin{cases} x(t) = x_0 + v t \cos \theta \\ y(t) = y_0 + v t \sin \theta \end{cases} \quad (6)$$

где x_0 и y_0 – координаты точки начала движения.

Таким образом, при использовании данной математической модели получена зависимость между заданным изменением декартовых координат точки на рабочей поверхности при перемещении манипулятора и полярных координат его звеньев. Это позволяет воспроизводить любой контур, заданный в виде отрезков прямой (полилинии) или набора произвольно расположенных контуров на рабочей поверхности.

При равномерном движении рабочей точки необходимо разбить отрезок прямой на множество равноотстоящих друг от друга точек, количеством которых будет определяться точность построения [2].

Основную погрешность в точности позиционирования механизма на основе двухзвенного манипулятора будут вносить конструктивные особенности, связанные с переменным шагом единичного перемещения. Рассмотрим влияние конструктивных особенностей на точность перемещения рабочего инструмента, обусловленных дискретным изменением углов $\alpha(t)$ и $\beta(t)$. Величина единичного перемещения может быть определена векторным уравнением:

$$\vec{L} = \overrightarrow{\Delta X(\alpha, \beta)} + \overrightarrow{\Delta Y(\alpha, \beta)}. \quad (7)$$

Зададимся величиной единичного изменения углов (i – шаг изменения угла $\beta(t)$, j – шаг изменения угла $\alpha(t)$) и определим выражение для модуля единичного шага перемещения:

$$L = \sqrt{(X_1(\alpha + j, \beta + i) - X(\alpha, \beta))^2 + (Y_1(\alpha + j, \beta + i) - Y(\alpha, \beta))^2}. \quad (8)$$

Переменные X, Y определены соотношениями (4), значения переменных X_1, Y_1 можно определить из уравнений:

$$\begin{aligned} X_1(t) &= R \cos(\beta(t) + i) \sin(\alpha(t) + j) + R \sin(\beta(t) + i) \cos(\alpha(t) + j) + R \cos(\alpha(t) + j) \\ Y_1(t) &= R \sin(\beta(t) + i) \sin(\alpha(t) + j) - R \cos(\beta(t) + i) \cos(\alpha(t) + j) + R \sin(\alpha(t) + j) \end{aligned} \quad (9)$$

График распределения величины единичного перемещения по рабочей поверхности механизма приведён на рисунке 2. Уравнение для определения L после подстановки в него выражений для переменных получается весьма громоздким и в данной работе не приведено. Однако в общем случае эта зависимость имеет вид:

$$L = k(\alpha, \beta) \cdot R,$$

где $k(\alpha, \beta)$ – коэффициент пропорциональности, зависящий от углов поворота диска и штанги.

Таким образом, величина погрешности прямопропорциональна радиусу диска и зависит сложным образом от текущей координаты рабочей точки.

На графике значение коэффициента ошибки $k(\alpha, \beta)$ приведено в безразмерных единицах. Дискретность изменения переменных α, β принята равной 0,0314 рад.

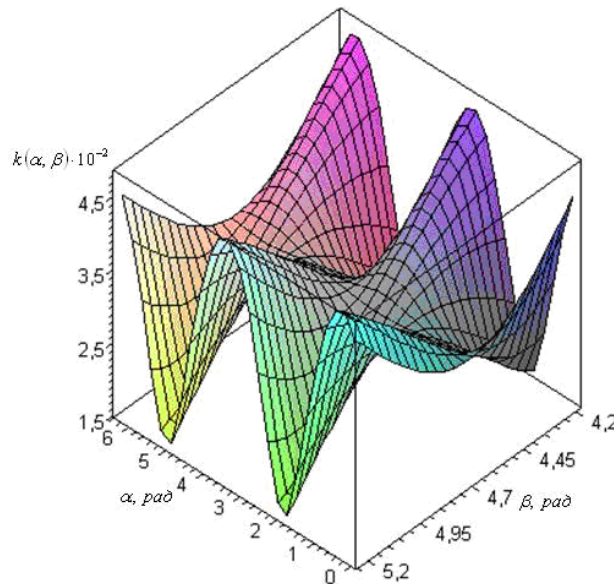


Рис. 2. Зависимость величины единичного смещения рабочей точки от угла поворота штанги при неподвижном диске

Разработан и изготовлен опытный образец станка с шаговым электроприводом и система управления. Проводится опытная эксплуатация устройства в качестве станка для раскроя и гравировки плоских материалов.

Список литературы

1. Емельянова И.В. Анализ и синтез приводов подач токарных станков с ЧПУ с целью повышения точности обработки: автореф. дис. ... канд. техн. наук: 05.02.08; 05.03.01 / Самарский гос. техн. ун-т. – Самара, 1995. – 23 с.
2. Теоретические основы САПР / В.П. Корячко, В.М. Курейчик, И.П. Норенков. – М.: Энергоатомиздат, 1987. – 400 с.
3. Кошкин В.Л. Аппаратные системы числового программного управления. – М.: Машиностроение, 1989. – 284 с.
4. Ключев А.С. Оптимизация автоматических систем управления по быстродействию. – М.: Энергоиздат, 1982.
5. Шевкопляс Б.В. Микропроцессорные структуры. Инженерные решения: справочник – 2-е изд., перераб. и доп. – М.: Радио и связь, 1990. – 512 с.: ил.

УДК: 621:004.9

О.Г. Толстунов

МОДЕЛИРОВАНИЕ ПРОСТРАНСТВЕННОГО МАНИПУЛЯТОРА С ГИБКИМ ПОДВЕСОМ ОБЪЕКТА ПЕРЕМЕЩЕНИЯ В СРЕДЕ MATLAB

Сегодня энергоэффективность и энергосбережение входят в 5 стратегических направлений приоритетного технологического развития, названных таковыми президентом РФ Дмитрием Медведевым на заседании Комиссии по модернизации и технологическому развитию экономики России, которое состоялось 18 июня 2009 г. Технологии энергосбережения должны быть использованы в различных отраслях, в том числе и на производстве. В связи с этим в данной работе предлагается пространственный манипулятор, который по сравнению с типовым подъёмно-транспортным оборудованием, предназначенным для перемещения объектов различного назначения в пространстве, обладает малой энерго- и материалоёмкостью, а также имеет расширенную зону обслуживания.

Целью данной работы является определение рабочей зоны обслуживания пространственного манипулятора для обеспечения безопасности процесса перемещения объектов.

Для достижения поставленной цели необходимо решить следующие основные задачи:

- определить зависимость длин гибких тросов от положения объекта в пространстве;
- определить зависимость сил натяжения в тросах пространственного манипулятора от положения объекта в пространстве;
- найти множество координат объекта перемещения, принадлежащих границе зоны обслуживания.

Для решения основных задач будет анализироваться кинематическая схема пространственного манипулятора, представленная на рисунке 1.

Предлагаемый пространственный манипулятор содержит четыре вертикальных колонны, установленные в точках $M_{11}(x_{11}; y_{11}; z_{11})$, $M_{22}(x_{22}; y_{22}; z_{22})$, $M_{33}(x_{33}; y_{33}; z_{33})$, $M_{44}(x_{44}; y_{44}; z_{44})$, в которых укреплены первые концы тросов. Вторые концы тросов соединяются в точке $P(x; y; z)$, перемещающейся в пространстве по заданной траектории. Перемещение объекта осуществляется посредством изменения длин гибких тросов L_1, L_2, L_3, L_4 .

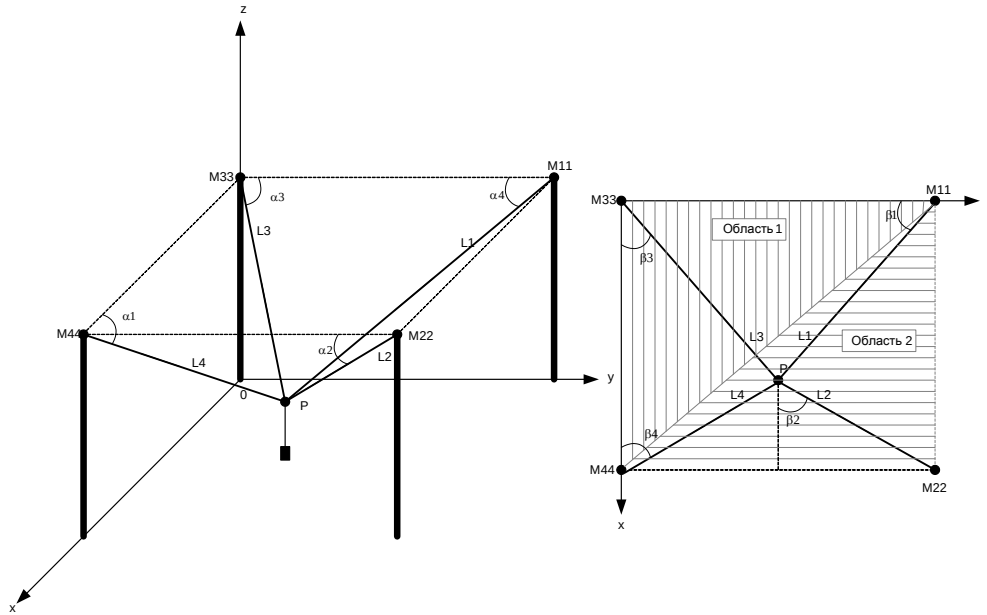


Рис. 1. Кинематическая схема пространственного манипулятора

Для нахождения зависимости длин тросов $L1$, $L2$, $L3$, $L4$ от положения точки $P(x; y; z)$ в пространстве найдём расстояние от точки $P(x; y; z)$ до каждой из точек – $M11(x11; y11; z11)$, $M22(x22; y22; z22)$, $M33(x33; y33; z33)$, $M44(x44; y44; z44)$:

$$\begin{cases} \overline{M11P} = L1(x, y, z) = \sqrt{(x11 - x)^2 + (y11 - y)^2 + (z11 - z)^2} \\ \overline{M22P} = L2(x, y, z) = \sqrt{(x22 - x)^2 + (y22 - y)^2 + (z22 - z)^2} \\ \overline{M33P} = L3(x, y, z) = \sqrt{(x33 - x)^2 + (y33 - y)^2 + (z33 - z)^2} \\ \overline{M44P} = L4(x, y, z) = \sqrt{(x44 - x)^2 + (y44 - y)^2 + (z44 - z)^2} \end{cases} \quad (1)$$

Как видно из рисунка 1, расстояние от точки $P(x; y; z)$ до каждой из точек $M11(x11; y11; z11)$, $M22(x22; y22; z22)$, $M33(x33; y33; z33)$, $M44(x44; y44; z44)$ является зависимостью длины тросов $L1$, $L2$, $L3$, $L4$ от положения точки $P(x; y; z)$ в пространстве.

При исследовании предлагаемого пространственного манипулятора с гибким подвесом объекта перемещения весьма важной задачей является расчёт сил натяжения, возникающих в тросах под действием силы тяжести объекта перемещения. На основании расчёта сил натяжения тросов пространственного манипулятора можно предъявить требования к прочности материала изготовления тросов, а также определить зону допустимых перемещений объекта.

На рис. 2 представлена кинематическая схема пространственного манипулятора с обозначенными на ней силами натяжения тросов и силой тяжести объекта перемещения.

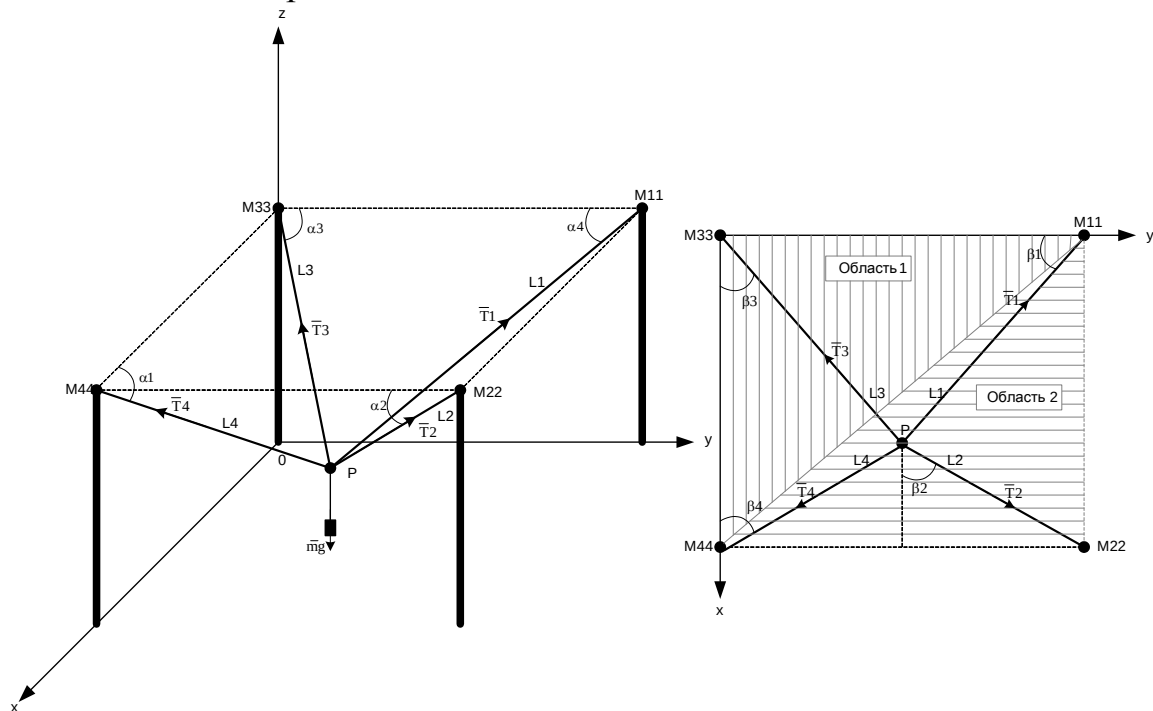


Рис. 2. Кинематическая схема пространственного манипулятора с обозначением сил натяжения тросов и силы тяжести объекта перемещения

При расчёте сил натяжения в тросах в зависимости от положения точки $P(x; y; z)$ в пространстве следует учитывать тот факт, что в действительности сила тяжести объекта перемещения, закреплённого в точке $P(x; y; z)$, распределяется только между тремя тросами, а силой натяжения четвёртого троса при этом можно пренебречь. Для выбора тросов, в которых вес объекта перемещения $P = m \cdot g$ создаёт ненулевые силы натяжения, разделим плоскость XOY на две равные треугольные области значений. При движении точки $P(x; y; z)$ в пределах значений из области 1 вес объекта перемещения $P = m \cdot g$ распределится между тросами, закреплёнными в точках $M11(x11; y11; z11)$, $M33(x33; y33; z33)$, $M44(x44; y44; z44)$, а при движении точки $P(x; y; z)$ в пределах значений из области 2 – между тросами, закреплёнными в точках $M11(x11; y11; z11)$, $M22(x22; y22; z22)$, $M44(x44; y44; z44)$.

Для определения принадлежности точки $P(x; y; z)$ той или иной области значений, на которые разделена зона обслуживания пространственного манипулятора, воспользуемся условием равенства площади одной из

областей значений сумме площадей других трёх вспомогательных треугольных областей значений, находящихся внутри рассматриваемой области.

Предположим, что точка $P(x; y; z)$ принадлежит первой области значений, образованной точками $M11(x11; y11; z11)$, $M33(x33; y33; z33)$ и $M44(x44; y44; z44)$, тогда в этой области возможно рассмотрение трёх вспомогательных треугольных областей значений, первая из которых заключена между точками $M11(x11; y11; z11)$, $P(x; y; z)$ и $M33(x33; y33; z33)$, вторая, вспомогательная, область значений заключена между точками $M33(x33; y33; z33)$, $P(x; y; z)$ и $M44(x44; y44; z44)$, третья – между точками $M11(x11; y11; z11)$, $P(x; y; z)$ и $M44(x44; y44; z44)$ (рис. 3).

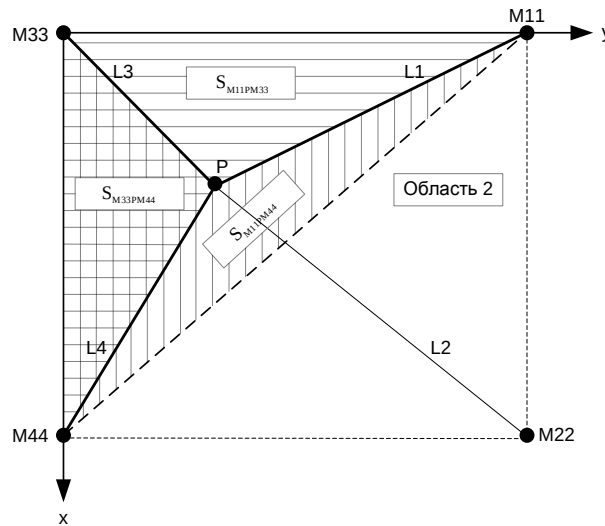


Рис. 3. Кинематическая схема пространственного манипулятора с гибким подвесом объекта перемещения в плоскости XOY при нахождении точки $P(x; y; z)$ в первой области значений

Найдём площадь области значений 1, воспользовавшись следующим выражением:

$$S_{M11M33M44} = \frac{1}{2} \begin{vmatrix} x11 & y11 & 1 \\ x33 & y33 & 1 \\ x44 & y44 & 1 \end{vmatrix}. \quad (2)$$

Площади вспомогательных областей можно рассчитать по формулам:

$$S_{M11PM33} = \frac{1}{2} \begin{vmatrix} x11 & y11 & 1 \\ x & y & 1 \\ x33 & y33 & 1 \end{vmatrix}, \quad (3)$$

$$S_{M33PM44} = \frac{1}{2} \begin{vmatrix} x_{33} & y_{33} & 1 \\ x & y & 1 \\ x_{44} & y_{44} & 1 \end{vmatrix}, \quad (4)$$

$$S_{M11PM44} = \frac{1}{2} \begin{vmatrix} x_{11} & y_{11} & 1 \\ x & y & 1 \\ x_{44} & y_{44} & 1 \end{vmatrix}. \quad (5)$$

Если $S_{M11M33M44} = S_{M11PM33} + S_{M33PM44} + S_{M11PM44}$, то точка $P(x; y; z)$ принадлежит области значений 1, в случае когда $S_{M11M33M44} \neq S_{M11PM33} + S_{M33PM44} + S_{M11PM44}$, точка $P(x; y; z)$ находится за пределами рассматриваемой области значений.

Если предположить, что точка $P(x; y; z)$ принадлежит второй области значений, образованной точками $M11(x_{11}; y_{11}; z_{11})$, $M22(x_{22}; y_{22}; z_{22})$ и $M44(x_{44}; y_{44}; z_{44})$, тогда в этой области возможно рассмотрение трёх вспомогательных треугольных областей значений, первая из которых заключена между точками $M11(x_{11}; y_{11}; z_{11})$, $P(x; y; z)$ и $M44(x_{44}; y_{44}; z_{44})$, вторая, вспомогательная, область значений заключена между точками $M22(x_{22}; y_{22}; z_{22})$, $P(x; y; z)$ и $M44(x_{44}; y_{44}; z_{44})$, третья – между точками $M11(x_{11}; y_{11}; z_{11})$, $P(x; y; z)$ и $M22(x_{22}; y_{22}; z_{22})$ (рис. 4).

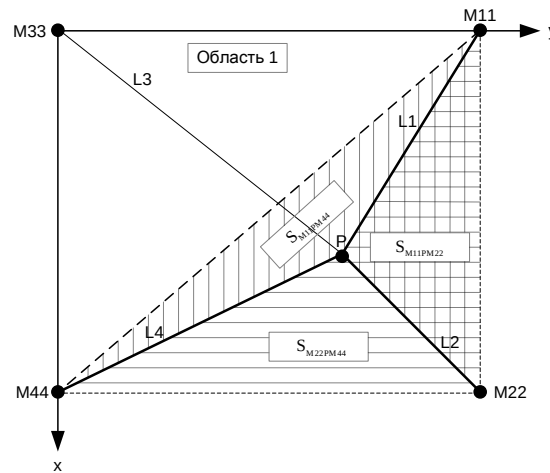


Рис. 4. Кинематическая схема пространственного манипулятора с гибким подвесом объекта перемещения в плоскости XOY при нахождении точки $P(x; y; z)$ во второй области значений

Найдём площадь области значений 2, воспользовавшись следующим выражением:

$$S_{M11M22M44} = \frac{1}{2} \begin{vmatrix} x_{11} & y_{11} & 1 \\ x_{22} & y_{22} & 1 \\ x_{44} & y_{44} & 1 \end{vmatrix}. \quad (6)$$

Площади вспомогательных областей можно рассчитать по формулам:

$$S_{M11PM44} = \frac{1}{2} \begin{vmatrix} x_{11} & y_{11} & 1 \\ x & y & 1 \\ x_{44} & y_{44} & 1 \end{vmatrix}, \quad (7)$$

$$S_{M22PM44} = \frac{1}{2} \begin{vmatrix} x_{22} & y_{22} & 1 \\ x & y & 1 \\ x_{44} & y_{44} & 1 \end{vmatrix}, \quad (8)$$

$$S_{M11PM22} = \frac{1}{2} \begin{vmatrix} x_{11} & y_{11} & 1 \\ x & y & 1 \\ x_{22} & y_{22} & 1 \end{vmatrix}. \quad (9)$$

Если $S_{M11M22M44} = S_{M11PM44} + S_{M22PM44} + S_{M11PM22}$, то точка $P(x; y; z)$ принадлежит области значений 2, в случае когда $S_{M11M22M44} \neq S_{M11PM44} + S_{M22PM44} + S_{M11PM22}$, точка $P(x; y; z)$ находится за пределами рассматриваемой области значений.

Если точка $P(x; y; z)$ принадлежит первой области значений, то на основании баланса сил можно найти силы натяжения тросов, закреплённых в точках $M11(x_{11}; y_{11}; z_{11})$, $M33(x_{33}; y_{33}; z_{33})$, $M44(x_{44}; y_{44}; z_{44})$, решив следующую систему уравнений:

$$\begin{cases} T1 \cdot \sin \alpha_1 + T3 \cdot \sin \alpha_1 + T4 \cdot \sin \alpha_4 = m \cdot g \\ -T1 \cdot \cos \alpha_0 \cdot \sin \beta_1 - T3 \cdot \cos \alpha_0 \cdot \cos \beta_0 + T4 \cdot \cos \alpha_0 \cdot \cos \beta_0 = 0, \\ T1 \cdot \cos \alpha_0 \cdot \cos \beta_0 - T3 \cdot \cos \alpha_0 \cdot \sin \beta_1 - T4 \cdot \cos \alpha_0 \cdot \sin \beta_1 = 0 \end{cases} \quad (10)$$

где $T1$ – сила натяжения троса, конец которого закреплён в точке $M11(x_{11}; y_{11}; z_{11})$; $T3$ – сила натяжения троса, конец которого закреплён в точке $M33(x_{33}; y_{33}; z_{33})$; $T4$ – сила натяжения троса, конец которого закреплён в точке $M44(x_{44}; y_{44}; z_{44})$. Значение α_1 является углом наклона троса, закреплённого в точке $M11(x_{11}; y_{11}; z_{11})$, к плоскости XOY при $Z = z_{11} = z_{22} = z_{33} = z_{44}$.

Значение угла α_1 можно найти из соотношения:

$$\alpha_1 = \arcsin \frac{z_{11} - z}{L1(x, y, z)}, \quad (11)$$

где z_{11} – высота первой вертикальной опоры; z – текущая координата точки P ; $L1(x, y, z)$ – длина троса, закреплённого в точке $M11(x_{11}; y_{11}; z_{11})$.

Значение α_3 представляет собой угол наклона троса, закреплённого в точке $M_{33}(x_{33}; y_{33}; z_{33})$, к плоскости XOY при $Z = z_{11} = z_{22} = z_{33} = z_{44}$. Значение угла α_3 можно найти из соотношения:

$$\alpha_3 = \arcsin \frac{z_{33} - z}{L_3(x, y, z)}, \quad (12)$$

где z_{33} – высота третьей вертикальной опоры; z – текущая координата точки P , $L_3(x, y, z)$ – длина троса, закреплённого в точке $M_{33}(x_{33}; y_{33}; z_{33})$.

Значение α_4 является углом наклона троса, закреплённого в точке $M_{44}(x_{44}; y_{44}; z_{44})$, к плоскости XOY при $Z = z_{11} = z_{22} = z_{33} = z_{44}$. Значение угла α_4 можно найти из соотношения:

$$\alpha_4 = \arcsin \frac{z_{44} - z}{L_4(x, y, z)}, \quad (13)$$

где z_{44} – высота четвёртой вертикальной опоры; z – текущая координата точки P ; $L_4(x, y, z)$ – длина троса, закреплённого в точке $M_{44}(x_{44}; y_{44}; z_{44})$.

Значение β_1 является углом наклона проекции троса, закреплённого в точке $M_{11}(x_{11}; y_{11}; z_{11})$, к оси OY . Значение угла β_1 можно найти из выражения:

$$\beta_1 = \arcsin \frac{x - x_{11}}{\sqrt{(x - x_{11})^2 + (y_{11} - y)^2}}, \quad (14)$$

где x – текущая координата точки P ; x_{11} – координата первой вертикальной опоры; $\sqrt{(x - x_{11})^2 + (y_{11} - y)^2}$ – длина троса, спроецированного на плоскость XOY , закреплённого в точке $M_{11}(x_{11}; y_{11}; z_{11})$.

Значение β_3 является углом наклона проекции троса, закреплённого в точке $M_{33}(x_{33}; y_{33}; z_{33})$, к оси OX . Значение угла β_3 можно найти из выражения:

$$\beta_3 = \arcsin \frac{y - y_{33}}{\sqrt{(x - x_{33})^2 + (y - y_{33})^2}}, \quad (15)$$

где y – текущая координата точки P ; y_{33} – координата третьей вертикальной опоры; $\sqrt{(x - x_{33})^2 + (y - y_{33})^2}$ – длина троса, спроецированного на плоскость XOY , закреплённого в точке $M_{33}(x_{33}; y_{33}; z_{33})$.

Значение β_4 является углом наклона проекции троса, закреплённого в точке $M_{33}(x_{33}; y_{33}; z_{33})$, к оси OX . Значение угла β_4 можно найти из выражения:

$$\beta_4 = \arcsin \frac{y - y_{44}}{\sqrt{(x_{44} - x)^2 + (y - y_{44})^2}}, \quad (16)$$

где y – текущая координата точки P ; y_{44} – координата четвёртой вертикальной опоры; $\sqrt{(x_{44} - x)^2 + (y - y_{44})^2}$ – длина троса, спроецированного на плоскость XOY , закреплённого в точке $M_{44}(x_{44}; y_{44}; z_{44})$.

В случае когда точка $P(x; y; z)$ принадлежит второй области значений, на основании баланса сил можно найти силы натяжения тросов, закреплённых в точках $M_{11}(x_{11}; y_{11}; z_{11})$, $M_{22}(x_{22}; y_{22}; z_{22})$, $M_{44}(x_{44}; y_{44}; z_{44})$, решив следующую систему уравнений:

$$\begin{cases} T_1 \cdot \sin \alpha_1 + T_2 \cdot \sin \alpha_1 + T_4 \cdot \sin \alpha_1 = m \cdot g \\ -T_1 \cdot \cos \alpha_0 \cdot \sin \beta_1 + T_2 \cdot \cos \alpha_0 \cdot \cos \beta_0 + T_4 \cdot \cos \alpha_0 \cdot \cos \beta_0 = 0, \\ T_1 \cdot \cos \alpha_0 \cdot \cos \beta_0 + T_2 \cdot \cos \alpha_0 \cdot \cos \beta_0 - T_4 \cdot \cos \alpha_0 \cdot \sin \beta_1 = 0 \end{cases} \quad (17)$$

где T_1 – сила натяжения троса, конец которого закреплён в точке $M_{11}(x_{11}; y_{11}; z_{11})$; T_2 – сила натяжения троса, конец которого закреплён в точке $M_{22}(x_{22}; y_{22}; z_{22})$; T_4 – сила натяжения троса, конец которого закреплён в точке $M_{44}(x_{44}; y_{44}; z_{44})$.

Значения углов α_1 , α_4 , β_1 , β_4 соответственно рассчитываются из выражений (17), (19), (20), (22). Значение α_2 является углом наклона троса, закреплённого в точке $M_{22}(x_{22}; y_{22}; z_{22})$, к плоскости XOY при $Z = z_{11} = z_{22} = z_{33} = z_{44}$. Значение угла α_2 можно найти из соотношения:

$$\alpha_2 = \arcsin \frac{z_{22} - z}{L_2(x, y, z)}, \quad (18)$$

где z_{22} – высота третьей вертикальной опоры; z – текущая координата точки P , $L_2(x, y, z)$ – длина троса, закреплённого в точке $M_{22}(x_{22}; y_{22}; z_{22})$.

Значение β_2 является углом наклона проекции троса, закреплённого в точке $M_{22}(x_{22}; y_{22}; z_{22})$, к оси OX . Значение угла β_2 можно найти из выражения:

$$\beta_2 = \arcsin \frac{y_{22} - y}{\sqrt{(x_{22} - x)^2 + (y_{22} - y)^2}}, \quad (19)$$

где y – текущая координата точки P ; y_{22} – координата второй вертикальной опоры; $\sqrt{(x_{22} - x)^2 + (y_{22} - y)^2}$ – длина троса, спроецированного на плоскость XOY , закреплённого в точке $M_{22}(x_{22}; y_{22}; z_{22})$.

На основании полученных в данной статье математических выражений в среде MATLAB была реализована компьютерная модель, позволяющая рассчитать множество координат объекта перемещения, при которых сила натяжения в любом из тросов превышает силу тяжести груза. Результат моделирования такого процесса представлен на рисунке 5.

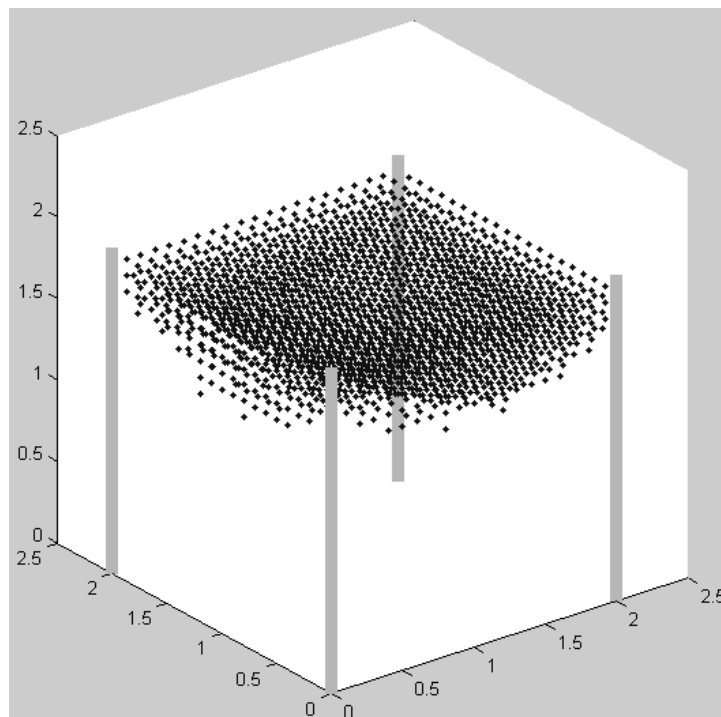


Рис. 5. Результат моделирования в среде MATLAB

При анализе полученных систем уравнений (10) и (17), описывающих статический процесс распределения силы тяжести объекта перемещения между тросами пространственного манипулятора, был замечен эффект увеличения сил натяжения при уменьшении значений углов $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, и в пределе, когда $\alpha_1 \rightarrow 0, \alpha_2 \rightarrow 0, \alpha_3 \rightarrow 0, \alpha_4 \rightarrow 0$, силы натяжения тросов пространственного манипулятора будут принимать бесконечно большие значения, то есть $T_1 \rightarrow \infty, T_2 \rightarrow \infty, T_3 \rightarrow \infty, T_4 \rightarrow \infty$. Замеченный эффект является недостатком предлагаемого пространственного манипулятора с гибким подвесом объекта перемещения, так как он накладывает ограничения на зону обслуживания манипулятора. Данный недостаток будет в значительной степени проявляться при практической реализации

пространственного манипулятора по причине конечного момента силы вращения оси электроприводов и ограниченной прочности механической конструкции манипулятора в целом.

Список литературы

1. Толстунов О.Г., Алепко А.В., Наумов И.И., Зеленский А.А. Устройство перемещения грузов // Актуальные проблемы техники и технологии: сб. тр. Всерос. науч.-практич. конф. – Шахты: ЮРГУЭС, 2009.
2. Толстунов О.Г., Валюкевич Ю.А. Пространственный манипулятор с гибким подвесом объекта перемещения // сб. тр. IV Всерос. межвуз. конф. молодых учёных. – СПб.: СПбГУ ИТМО, 2009. – С. 247–253.
3. Толстунов О.Г., Валюкевич Ю.А. Пространственный манипулятор на основе гибких механических связей // Мехатроника, автоматизация, управление (МАУ – 2009): сб. тр. Междунар. науч.-технич. конф. – Таганрог: ТТИ ЮФУ, 2009. – С. 314–316.
4. Толстунов О.Г., Валюкевич Ю.А. Моделирование кинематики тросового грузоподъёмного механизма в среде MATLAB // Информационные системы и технологии. Теория и практика: сб. науч. тр. – Шахты: ГОУ ВПО «ЮРГУЭС», 2009. – С. 122–127.

УДК 621.01:004.9

А.В. Алепко

МОДЕЛЬ ТРОСА С КВАЗИРАСПРЕДЕЛЁННЫМИ ПАРАМЕТРАМИ ДЛЯ ПРОСТРАНСТВЕННОГО МАНИПУЛЯТОРА С ГИБКИМ ПОДВЕСОМ ОБЪЕКТА ПЕРЕМЕЩЕНИЯ

Традиционно для перемещения грузов при складировании используются краны и манипуляторы, представляющие собой конструкцию с жёстко связанными элементами, которая имеет большую массу и перемещается по рельсовому пути, занимая при этом достаточно много полезной площади. Альтернативой таким конструкциям могут служить манипуляторы на основе гибких связей (рис. 1), в которых роль жёстких звеньев выполняют тросовые системы [1]. Однако для управления такими устройствами необходимо учитывать упругие свойства тросов, которые будут проявляться в колебании длины троса за счёт растяжения и сжатия под действием как гравитационных сил, так и использующихся приводов. Оценить деформацию тросов возможно путём создания математической модели, учитывающей конструктивные особенности исследуемого манипулятора и массу груза.

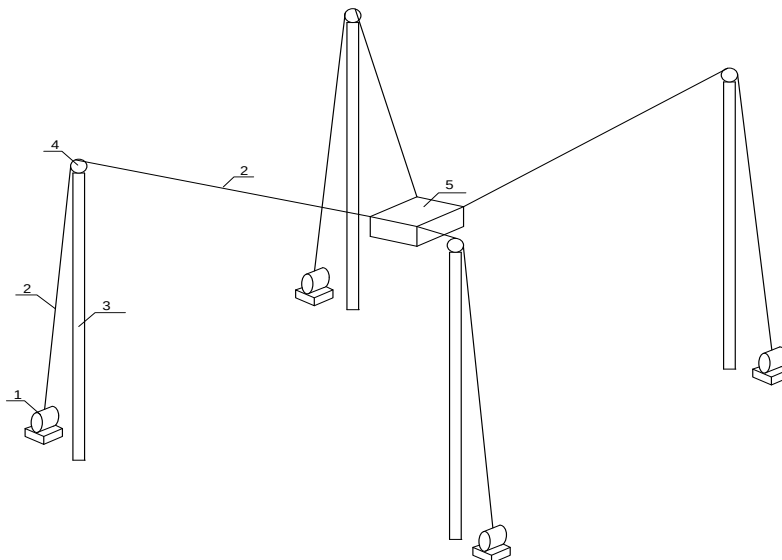


Рис. 1. Пространственный манипулятор с гибким подвесом объекта перемещения

Рассмотрим модель троса с квазираспределёнными параметрами на примере манипулятора на рисунке 1. Поскольку тросовые системы колонн идентичны, то для построения модели целесообразно рассмотреть одну такую систему (рис. 2).

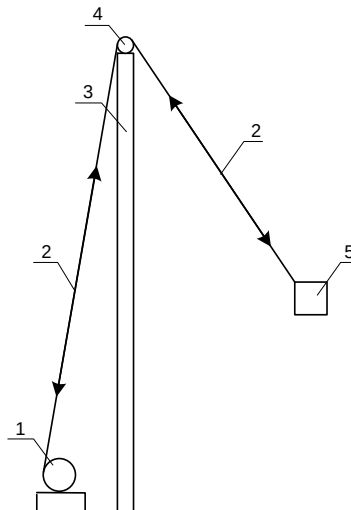


Рис. 2. Тросовая система колонны

На рисунках 1 и 2 приняты следующие обозначения:

- 1 – барабан для намотки троса с управляемым приводом;
- 2 – трос;
- 3 – вертикальная колонна;
- 4 – подвижный ролик;
- 5 – перемещаемый груз.

Разобьём структуру на три независимых элемента (рис. 3), поскольку на данных участках на трос действуют различные совокупности сил – участок троса между приводом и подвижным роликом, участок троса на ролике и участок троса между роликом и грузом.

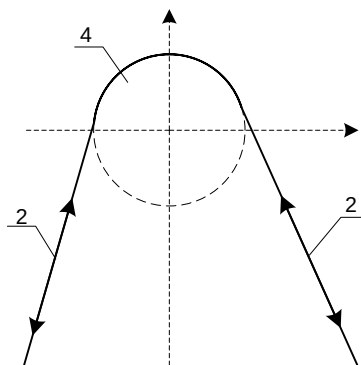


Рис. 3. Движение троса через подвижный ролик

Под квазираспределёнными характеристиками примем разбиение троса на набор сосредоточенных масс, соединённых между собой упругими связями, деформируемыми по закону Гука (рис. 4) [2]. Для тонкого растяжимого стержня, которым по сути является участок троса малой длины, закон имеет вид:

$$F = -k\Delta l, \quad (1)$$

где F – сила упругости участка троса; k – коэффициент жёсткости; Δl – удлинение участка троса. Одновременно с этим F – сила, действующая на центр сосредоточенной массы в предлагаемой модели: $F = m_g g$, где m_g – масса участка троса до текущей точки и масса груза с учётом угла наклона троса.

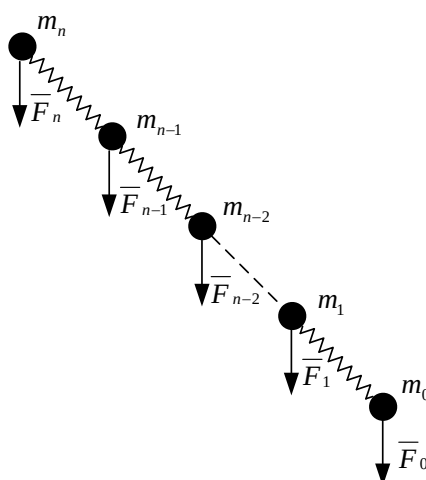


Рис. 4. Квазираспределённая модель троса

В связи с тем что при намотке троса на барабан он остаётся деформированным, необходимо оперировать изменением всей максимально возможной рабочей длины троса.

На основе данной модели в среде Codegear C++ Builder была создана компьютерная модель, визуализирующая изменение длины троса по координатам (рис. 5).

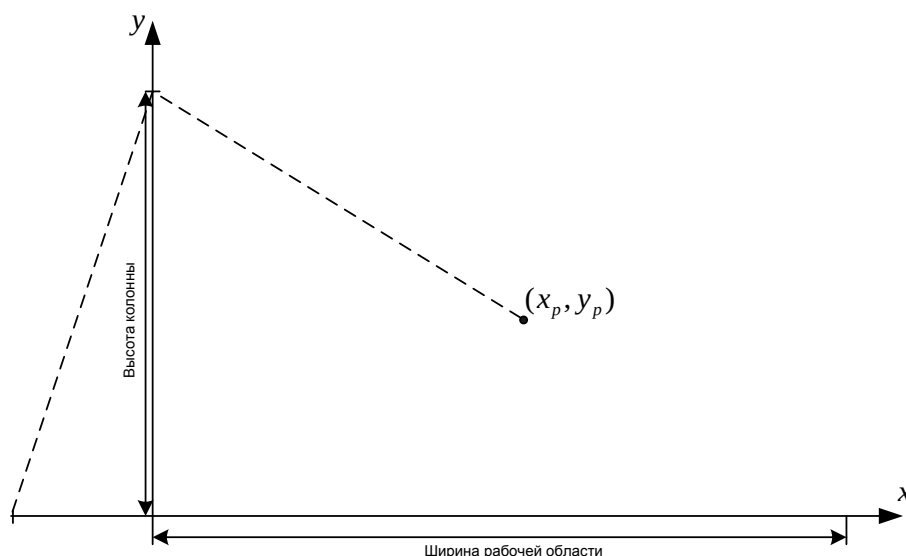


Рис. 5. Упрощённая геометрическая модель

Для упрощения модели участок троса, на котором происходит движение через ролик, был заменён эквивалентной передаточной функцией, не учитывающей его геометрические размеры. Коэффициент жёсткости троса был принят равным $5 \cdot 10^5$, высота колонны – 500, максимальная ширина рабочей области – 1000. Шаг разбиения был выбран равным 5. Масса груза – 10. Результаты моделирования представлены на рисунке 6.

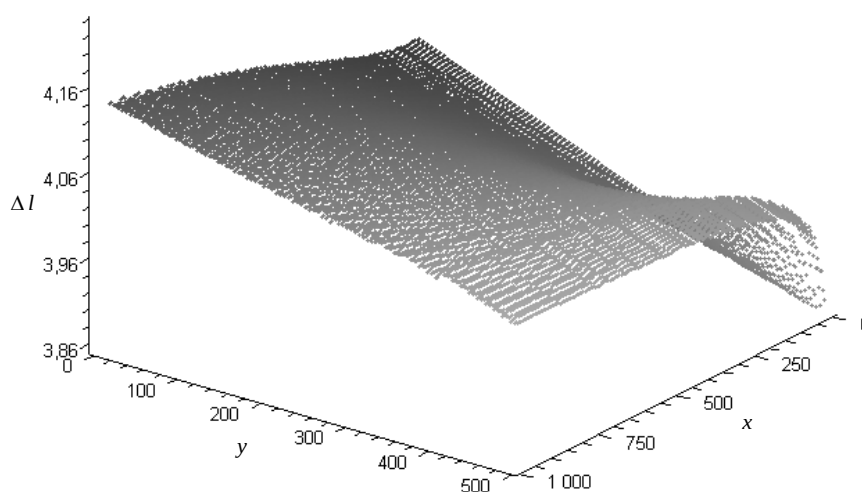


Рис. 6. Относительное удлинение троса в зависимости от расположения груза

На основании полученных результатов видно, что помимо груза значительное влияние на деформацию оказывает собственная масса троса. Разработанная модель позволяет учитывать данные факторы и при дальнейшем развитии в качестве компонента для системы управления манипулятором способствовать их нейтрализации.

Список литературы

1. Толстунов О.Г., Алепо А.В., Наумов И.И., Зеленский А.А. Устройство перемещения грузов // Актуальные проблемы техники и технологии: сб. тр. Всерос. науч.-практич. конф. – Шахты: ЮРГУЭС, 2009.
2. Заболотнов Ю.М., Фефелов Д.И. Математическая модель движения тросовой системы с распределёнными параметрами // Труды Всерос. науч. конф. (26–28 мая 2004 г.). Ч. 2. Моделирование и оптимизация динамических систем и систем с распределёнными параметрами, Матем. моделирование и краевые задачи. – Самара: СамГТУ, 2004. – С. 86–88.

УДК 54:620.179.17

П.Н. Козаченко, В.В. Дубовсков

ПРОГРАММНО-АППАРАТНЫЙ КОМПЛЕКС ЦИФРОВОЙ ОБРАБОТКИ АКУСТИКО-ЭМИССИОННЫХ СИГНАЛОВ ХИМИЧЕСКИХ РЕАКЦИЙ

Наряду с широко распространёнными и зарекомендовавшими себя физическими методами изучения кинетики протекания химических реакций, в последнее время в научных изданиях [1] выдвигаются предположения об использовании методов акустической эмиссии для анализа химических процессов. Физическая природа возникновения акустической эмиссии (АЭ) связана с микропроцессом деформирования и разрушения твёрдых материалов [2]. Применительно к твёрдым материалам этот метод получил широкое распространение. В первую очередь в дефектоскопии, где целью АЭ контроля обычно является обнаружение, определение координат и слежение (мониторинг) за источниками акустической эмиссии, связанными с дефектами металла, объектов или сварных соединений, а также для оценки скорости развития дефекта в целях заблаговременного прекращения эксплуатации или испытаний и предотвращения разрушения изделия. Именно поэтому сравнительно редко публикуются материалы, посвящённые использованию этого метода для изучения процессов в жидкости [3]; и в большинстве своём природа акустической эмиссии в жидкости обычно связывается с генерацией акустических волн твёрдыми веществами, например, при процессах плавления или кристаллизации [1]. В то же время

диагностика физико-химических процессов в жидкости также нуждается в разработке бесконтактного и исключительно информативного метода, каким является метод АЭ.

Сигналы акустической эмиссии фиксируются в виде колебаний, при которых смещение поверхности образца составляет 10^{-10} – 10^{-7} м. Частотный спектр акустической эмиссии в твёрдых материалах весьма широк и простирается от слышимых частот до десятков и сотен МГц. Иногда эти сигналы достаточно сильны и могут восприниматься на слух. В то же время чувствительность и возможности АЭ длительное время не позволяли использовать этот метод для изучения тонких эффектов, связанных со структурными превращениями в жидкости. С развитием техники эксперимента и применением компьютерных технологий возможности метода значительно возросли.

Эффект возникновения акустических колебаний в гетерогенных реакциях и при фазовых переходах известен уже довольно давно. Например, в работе [1] обнаружена генерация акустических колебаний при кристаллизации воды, бензола и гипосульфита. Процессы плавления и кристаллизации этих веществ сопровождаются возникновением импульсов, лежащих в интервале частот 0,5–2000 кГц. Отмечено, что частота повторения звуковых сигналов связана со степенью переохлаждения системы.

Согласно представлениям классической химии, изменение свойств системы при протекании химических процессов происходит непрерывно и сопровождается выделением или поглощением тепловой энергии, количество которой определяется природой реагирующих веществ, и их физическим состоянием. Реакции осуществляются только при столкновении молекул реагентов, и, согласно классическим представлениям, имеет место непрерывное изменение концентрации реагирующих веществ и, следовательно, происходит монотонное изменение физических параметров системы. Обычно при изучении химических и физико-химических процессов исследуются тепловые эффекты. Однако принципиальных запретов на то, чтобы энергия в ходе химических реакций и физико-химических процессов выделялась в виде механического движения, электромагнитного излучения и других форм, не существует. Поскольку химические реакции связаны с изменением геометрии молекул, изменением степени упорядоченности вещества, а иногда сопровождаются и фазовыми переходами, в реакционной среде неизбежно должны генерироваться акустические сигналы. Так как любые реальные реагирующие системы (жидкости, стёкла, кристаллы) имеют определённую внутреннюю структуру [3], то при протекании химических реакций и физико-химических процессов последняя неизбежно должна меняться. Причём этот переход происходит в небольших объёмах вещества и должен сопровождаться скачками в изменении физических параметров. Таким образом, вместо непрерывных кинетиче-

ских кривых, которыми традиционно описываются реакционные системы, на самом деле любой химический и физико-химический процесс должен характеризоваться ступенчатым изменением температуры, концентрации и других параметров. Величина этих ступенек (скачков) чрезвычайно мала и определяется надмолекулярной структурой вещества. Изменение состояния системы на одну ступеньку обязательно должно сопровождаться акустическими колебаниями (излучением волн напряжений, акустической эмиссией).

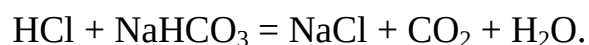
Явление генерации акустических колебаний в химических реакциях и физико-химических процессах, как прикладное, так и для проведения научных исследований, может найти широкое применение. Генерация акустических колебаний при протекании химических реакций и физико-химических процессов может успешно использоваться для исследования кинетики процессов, поскольку количество акустических импульсов в единицу времени характеризует скорость реакции. Данное явление может найти применение для создания приборов контроля за ходом реакции, при этом отпадает необходимость отбирать пробы реакционной среды для анализа.

Целью работы является применение методов обработки сигналов АЭ для анализа кинетики химических реакций.

Поскольку любая химическая реакция сопровождается выделением или поглощением энергии, это должно приводить к возникновению волн упругости (напряжения) в реакционной среде, которые можно обнаружить, используя датчики, обеспечивающие приём слабых сигналов на фоне собственных шумов. К возникновению акустических сигналов в химических реакциях должно приводить также изменение объёма продуктов реакции по сравнению с объёмом субстратов.

Для экспериментального подтверждения выдвинутых предположений использовался программно-аппаратный комплекс на базе ЭВМ. Структурная схема экспериментальной установки состоит из штатива с закреплённым узконаправленным микрофоном, колбы с реагирующими веществами, а также ЭВМ с платой сбора данных.

К системам с достаточно высоким уровнем акустической эмиссии можно отнести, например, гетерогенные реакции взаимодействия металлов, металлосодержащих смесей, солей и кислоты. Реакции взаимодействия металлов и кислот широко используются в промышленности. Исследование кинетики таких реакций контактными способами затруднительно в связи с наличием агрессивной среды. Для проведения эксперимента была выбрана реакция взаимодействия соды и соляной кислоты различной концентрации:



В ходе реакции выделяется углекислый газ, схлопывание пузырьков газа создаёт упругие волны, регистрируемые микрофоном. На основе спектрального и корреляционного анализа записанных сигналов было выявлено, что сигнал представляет собой случайный процесс, нестационарный по дисперсии. Спектральная плотность мощности записанных сигналов имеет схожую структуру и одинаковую полосу частот, что подтверждает химическую идентичность реагирующих веществ. Автокорреляционная функция представляет собой дельта-функцию. Этот фактор указывает на то, что исследуемые сигналы носят случайный характер. Как показывает анализ представленных рисунков, различия в концентрации кислоты отражаются в продолжительности процесса, при этом форма огибающей сохраняется, т.е. меняется её масштаб. В этом случае информационным параметром может служить огибающая процесса. Огибающая нестационарного по дисперсии случайного процесса представляет собой функцию изменения амплитуды пиковых значений процесса во времени.

Кроме того, можно сделать предположение, что огибающая процесса количественно отражает образование газа в ходе реакции, т.к. именно образовавшиеся пузырьки газа при схлопывании создают упругие волны, воздействующие на микрофон.

Известные методы выделения огибающей основаны на использовании пикового детектора, спектральных методов преобразования Гильберта и мало подходят для выделения функции нестационарности по дисперсии случайного процесса.

Для выделения огибающей сигналов акустической эмиссии предлагается метод, основанный на сравнении в скользящем окне отсчётов реализации. Первый отсчёт сравнивается с последующим. В случае если последующий отсчёт меньше предыдущего, они приравняются. При невыполнении условия $Y_i < Y_{i+1}$ окно сдвигается на один отсчёт. При достижении скользящим окном конца реализации вновь полученная реализация сглаживается с помощью метода кусочного разложения оценок [4].

На рисунке 1а представлены сигнал акустической эмиссии рассматриваемой ранее реакции и огибающая, выделенная с помощью предлагаемого метода, а на рисунке 1б – огибающие рассматриваемой реакции при различных концентрациях кислоты.

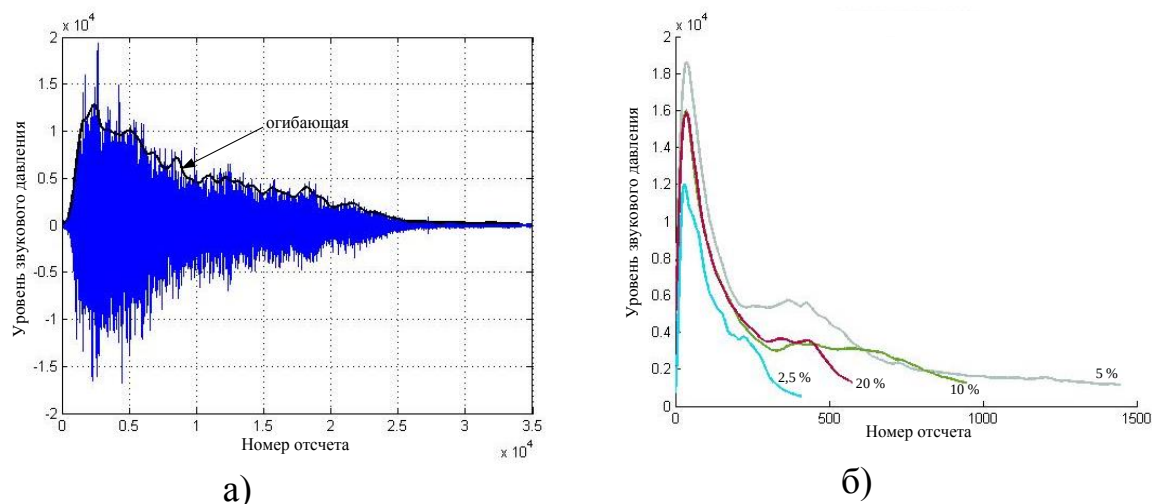


Рис. 1. Сигнал акустической эмиссии и выделенная огибающая, реакции взаимодействия соды и соляной кислоты при концентрации 2,5 % (а), огибающие при различной концентрации кислоты (б)

В результате выполнения работы получены научно-технические результаты:

- проведённый анализ существующих методов исследования и контроля химических реакций показал, что область их использования в условиях промышленных и технологических циклов производств химической отрасли затруднён сложностью и дороговизной используемого оборудования;
- разработан способ анализа химических реакций, основанный на использовании методов акустической эмиссии;
- на основе разработанного метода предлагается программный комплекс обработки сигналов акустической эмиссии.

Список литературы

1. Смирнов А.Н. Генерация акустических колебаний в химических реакциях и физико-химических процессах // Рос. хим. ж. – 2001. – Т. 45. – С. 29–34.
2. Грешников В.А., Дробот Ю.Б. Акустическая эмиссия. – М.: Изд-во Стандартов, 1976.
3. Смирнов А.Н., Дементьев А.Н. // Журнал физической химии. – 1985. – Т. 59, вып. 7. – С. 1719–1721.
4. Марчук В.И. Первичная обработка результатов измерений при ограниченном объёме априорной информации: монография / под ред. К.Е. Румянцева. – Таганрог: Изд-во ТРТУ, 2003. – 160 с.

УДК 519.876.5

Д.А. Безуглов, И.В. Андриющенко, С.А. Швидченко

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ИДЕНТИФИКАЦИИ ЗАКОНА РАСПРЕДЕЛЕНИЯ НА БАЗЕ КУМУЛЯНТНОГО МЕТОДА АНАЛИЗА РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ

Разработан новый кумулянтный метод определения вида плотности распределения, а на его базе – информационная технология, позволяющая полностью автоматизировать процесс определения вида закона распределения результатов измерений. Правильность гипотезы может быть проверена в дальнейшем с помощью известных критериев, например Колмогорова или Стьюдента. Приведён пример реализации информационной технологии.

Ключевые слова: информационная технология идентификации вида плотности распределения; кумулянтный метод анализа статистических данных; моменты распределения; кумулянты распределения.

Введение. Современные тенденции развития компьютерных и информационных технологий в технике, в частности в метрологии, ориентированы на использование и разработку программных систем, обеспечивающих решение задач при любой форме регистрируемых измерений, включающих современные методы статистического анализа. Они ориентированы на широкое, но корректное использование в исследованиях методов компьютерного моделирования.

Под задачей идентификации закона распределения наблюдаемой случайной величины понимается нахождение такого закона, который бы в статистическом смысле соответствовал имеющимся наблюдениям. Случайные ошибки средств измерений не так уж часто подчиняются нормальному закону [1, 2]. В основе измерительных приборов и систем лежат различные физические принципы, различные методы измерений и различные преобразования измерительных сигналов. Погрешности измерений являются следствием влияния множества факторов случайного и неслучайного характера, действующих постоянно или эпизодически. Поэтому понятно, что только при выполнении определённых предпосылок (теоретических и технических) погрешности измерений достаточно хорошо описываются нормальным законом распределения.

При создании современных программных и аппаратных средств для автоматизированной обработки результатов измерений рассматривается задача идентификации закона распределения случайных величин, которая обычно решается с применением статистических критериев согласия Пирсона, Колмогорова, Колмогорова – Смирнова, омега-квадрат. В последние годы появились предложения об аппроксимации законов распределения без использования критериев согласия. Задача выбора закона распределе-

ния для данной выборки значений случайной величины решается в этом случае как задача приближения функции на заданных аналитических законах распределений. В качестве закона распределения выбирается такое распределение, которое находится на минимальном расстоянии (в качестве которого может быть выбрана статистика Пирсона, Колмогорова и др.) от эмпирического закона распределения, построенного по соответствующей выборке. Выбор числовых параметров законов распределений выполняется на основе обработки всех данных единственной выборки случайной величины, в то время как в использовании критериев согласия параметры требуется определять по иной выборке значений случайной величины.

Однако недостатком всех рассмотренных выше подходов является то, что гипотеза о виде распределения определяется эмпирически. То есть исследователь должен по виду полученной функции распределения вероятностей (интегральной или дифференциальной) исходя из своего опыта определить, каким законом она описывается.

Целью данной работы является разработка информационной технологии, позволяющей полностью автоматизировать процесс идентификации вида закона распределения на базе метода кумулянтного анализа результатов измерений. В дальнейшем правильность гипотезы может быть проверена с помощью известных критериев, например Колмогорова или Стьюдента.

1. Кумулянтный анализ

Кумулянтное описание случайных величин даёт столь же полное их статистическое представление, сколь и моментное. Оно обладает, вместе с тем, важными и привлекательными преимуществами [3]. Первое преимущество заключается в том, что кумулянты, в отличие от моментов, имеют чётко выраженный самостоятельный статистический смысл и могут быть заданы в определённой степени независимо друг от друга, являясь в этом плане некоторыми «нормальными координатами» статистического описания. Это приводит, например, к тому, что различные статистические средние «выходов» нелинейных преобразований выражаются простым образом именно через кумулянты «входных» переменных [3, 4].

Второе преимущество кумулянтов связано с тем, что учёт их высших порядков позволяет просто описать любую степень негауссовости случайных величин. По этой причине основную ценность кумулянтное описание имеет именно для негауссовых переменных [4].

Третье, весьма важное преимущество кумулянтного описания случайных величин обусловлено тем, что конечному набору кумулянтов всегда соответствует некоторая вещественная функция, аппроксимирующая вероятностное распределение. Это обстоятельство имеет особо важное значение при приближённом представлении вероятностных распределений тех случайных величин, для которых можно отыскать лишь конечные наборы кумулянтов [5].

2. Кумулянтный метод анализа результатов измерений

Рассмотрим сущность предлагаемого метода. Характеристическую функцию можно записать в виде $\Theta(u) = \exp(B(u))$, где, очевидно, должно быть $B(0) = 0$. Разложим функцию $B(u)$ в степенной ряд [3]:

$$\beta(u) = \ln \Theta(u) = \sum_{k=1}^{\infty} \frac{\chi_k}{k!} (ju)^k. \quad (1)$$

Коэффициенты этого ряда

$$\chi_k = j^{-k} \beta^{(k)}(0) = j^{-k} \left[\frac{d^k \ln \Theta(u)}{du^k} \right]_{u=0} \quad (2)$$

так же, как и моменты, являются характеристиками распределения и носят название кумулянтов или семи инвариантов. Таким образом, характеристическую функцию можно представить в виде:

$$\Theta(u) = \exp \left[\sum_{k=1}^{\infty} \frac{\chi_k}{k!} (ju)^k \right]. \quad (3)$$

Кумулянты однозначно определяют случайную величину, если ряд (3) сходится для всех u . Поэтому набор кумулянтов $\chi_1, \chi_2, \chi_3, \dots, \chi_s$ также может служить тождественным представлением закона распределения.

Если известны моменты α , то кумулянты могут быть найдены из следующих соотношений [3]:

$$\begin{aligned} \chi_1^i &= \alpha_1^i = m; & \chi_2^i &= \alpha_2^i - \alpha_1^{i2} = D; & \chi_3^i &= \alpha_3^i - 3\alpha_1^i \alpha_2^i + 2\alpha_1^{i3}; \\ \chi_4^i &= \alpha_4^i - 3\alpha_1^{i2} \alpha_2^i - 4\alpha_1^i \alpha_3^i + 12\alpha_1^{i2} \alpha_2^i - 6\alpha_1^{i4}; \\ \chi_5^i &= \alpha_5^i - 5\alpha_1^i \alpha_4^i - 10\alpha_2^i \alpha_3^i + 20\alpha_1^{i2} \alpha_3^i + 30\alpha_1^i \alpha_2^{i2} - 60\alpha_1^{i3} \alpha_2^i + 24\alpha_1^{i5}; \\ \chi_6^i &= \alpha_6^i - 6\alpha_1^i \alpha_5^i - 15\alpha_2^i \alpha_4^i + 30\alpha_1^{i2} \alpha_4^i - 10\alpha_3^{i2} + 120\alpha_1^i \alpha_2^i \alpha_3^i - \\ &- 120\alpha_1^{i3} \alpha_3^i + 30\alpha_2^{i3} - 270\alpha_1^{i2} \alpha_2^{i2} + 360\alpha_1^{i4} \alpha_2^i - 120\alpha_1^{i6}. \end{aligned} \quad (4)$$

Верхним индексом i в дальнейшем будем обозначать принадлежность моментов и кумулянтов к тому или иному распределению. В выражении (4) мы использовали только шесть первых кумулянтов. Это связано с тем, что, как показали проведённые авторами исследования, на различных выборках такого количества кумулянтов достаточно для раскрытия сущности предложенного подхода. Однако при создании программного продукта число кумулянтов может быть увеличено. При этом точность метода возрастёт.

В аналитическом виде моменты распределений могут быть вычислены через производящую функцию. Рассмотрим для примера выражения для моментов гауссовского, равномерного, экспоненциального и пуассоновского законов распределения. В дальнейшем эти выражения мы используем для подстановки в выражения (4) для определения кумулянтов соответствующих распределений.

Пусть имеется случайная величина X с распределением P^X . Тогда её производящей функцией моментов называется функция, имеющая вид:

$$M_X(t) = E[e^{tx}]. \quad (5)$$

Момент случайной величины вычисляется через производящую функцию следующим образом [4]:

$$E[X^n] = \frac{d^n}{dt^n} M_X(t). \quad (6)$$

Для распределения Гаусса

$$W_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \int \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dt \quad (7)$$

производящая функция моментов будет иметь вид [2]:

$$M_X(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right). \quad (8)$$

Соответственно, моменты могут быть вычислены в соответствии со следующими выражениями:

$$\begin{aligned} \alpha^2_1 = m = \mu; \quad \alpha^2_2 = \mu^2 + \sigma^2; \quad \alpha^2_3 = \mu^3 + 3\mu\sigma^2; \quad \alpha^2_4 = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4; \\ \alpha^2_5 = \mu^5 + 10\mu^3\sigma^2 + 15\mu\sigma^4; \quad \alpha^2_6 = \mu^6 + 15\mu^4\sigma^2 + 45\mu^2\sigma^4 + 15\sigma^6. \end{aligned} \quad (9)$$

Для равномерного распределения

$$W_p(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x < b, \\ 1, & x \geq b \end{cases} \quad (10)$$

производящая функция моментов будет иметь вид [4]:

$$E[X^n] = \frac{1}{n+1} \sum_{k=0}^n a^k b^{n-k}. \quad (11)$$

Соответственно, моменты могут быть найдены из следующих выражений:

$$\begin{aligned}\alpha^{p_1} = m &= \frac{1}{2}(a+b); \quad \alpha^{p_2} = \frac{1}{3}(a^2+ab+b^2); \quad \alpha^{p_3} = \frac{1}{4}(a^3+a^2b+ab^2+b^3); \\ \alpha^{p_4} &= \frac{1}{5}(a^4+a^3b+a^2b^2+ab^3+b^4); \quad \alpha^{p_5} = \frac{1}{6}(a^5+a^4b+a^3b^2+a^2b^3+ab^4+b^5); \\ \alpha^{p_6} &= \frac{1}{7}(a^6+a^5b+a^4b^2+a^3b^3+a^2b^4+ab^5+b^6).\end{aligned}\quad (12)$$

Для экспоненциального распределения

$$W_e(x) = \begin{cases} \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, & x > 0, \\ 0, & x < 0 \end{cases} \quad (13)$$

производящая функция моментов будет иметь вид [4]:

$$M_x(t) = (1 - \frac{t}{\lambda})^{-1}. \quad (14)$$

Соответственно, моменты могут быть найдены из следующих выражений:

$$\alpha^e_1 = m = \lambda^{-1}; \quad \alpha^e_2 = \frac{2}{\lambda^2}; \quad \alpha^e_3 = \frac{6}{\lambda^3}; \quad \alpha^e_4 = \frac{24}{\lambda^4}; \quad \alpha^e_5 = \frac{120}{\lambda^5}; \quad \alpha^e_6 = \frac{720}{\lambda^6}. \quad (15)$$

Для распределения Пуассона:

$$W_n(x) = \sum_{i=0}^x \frac{\lambda^i e^{-\lambda}}{i!} \quad (16)$$

производящая функция моментов будет иметь вид [4]:

$$M_x(t) = \exp(\lambda(e^t - 1)). \quad (17)$$

Соответственно, моменты могут быть найдены из следующих выражений:

$$\begin{aligned}\alpha^n_1 = m &= \lambda; \quad \alpha^n_2 = \lambda^2 + \lambda; \quad \alpha^n_3 = \lambda^3 + 3\lambda^2 + \lambda; \\ \alpha^n_4 &= \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda; \quad \alpha^n_5 = \lambda^5 + 10\lambda^4 + 25\lambda^3 + 15\lambda^2 + \lambda; \\ \alpha^n_6 &= \lambda^6 + 15\lambda^5 + 65\lambda^4 + 90\lambda^3 + 31\lambda^2 + \lambda.\end{aligned}\quad (18)$$

Для набора распределений вычисляются с использованием выражений (4), (9), (12), (18) теоретические значения кумулянтов. При этом при программной реализации предложенного подхода количество видов используемых распределений может быть значительно увеличено и ограничивается лишь рамками решаемой при этом метрологической задачи.

Пусть имеются результаты измерений Y_i . Эмпирические моменты α_i могут быть найдены на основе обработки входной реализации:

$$\alpha^{ei}_1 = \frac{\sum_{m=0}^{l-1} Y_m}{n}; \alpha^{ei}_2 = \frac{\sum_{m=0}^{l-1} Y_m^2}{n-1}; \alpha^{ei}_3 = \frac{l \sum_{m=0}^{l-1} Y_m^3}{(l-1)(l-2)}; \alpha^{ei}_4 = \frac{l^2 \sum_{m=0}^{l-1} Y_m^4}{(l-1)(l-2)(l-3)};$$

$$\alpha^{ei}_5 = \frac{l^3 \sum_{m=0}^{l-1} Y_m^5}{(l-1)(l-2)(l-3)(l-4)}; \alpha^{ei}_6 = \frac{l^4 \sum_{m=0}^{l-1} Y_m^6}{(l-1)(l-2)(l-3)(l-4)(l-5)}, \quad (19)$$

где l – размерность выборки; α^{ei}_j – эмпирический момент j -го порядка i -го закона распределения.

По ним в соответствии с выражением (19) определяются эмпирические значения кумулянтов. При этом необходимо обеспечить нормировку по второму кумулянту.

В данном случае критерием «близости» исследуемого распределения к тому или иному виду может служить следующее выражение:

$$S_i = \min_i \sum_{j=1}^6 (\chi_j^i - \chi_j^{ei})^2, \quad (20)$$

где i – вид распределения, $i=1, 2, 3$; χ_j^i – кумулянты, рассчитанные аналитически; χ_j^{ei} – кумулянты, рассчитанные эмпирически.

Предлагаемая информационная технология может быть реализована программно или аппаратно, как это показано на рисунке 1.

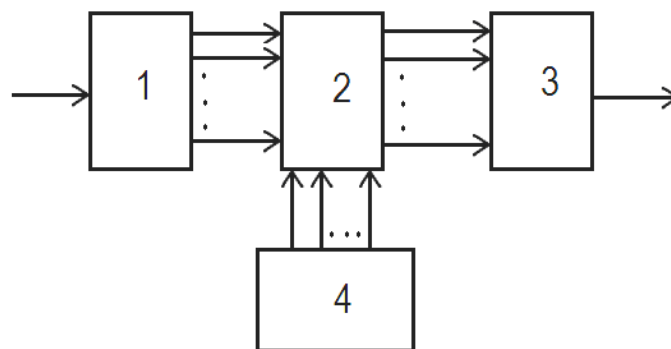


Рис. 1. Информационная технология идентификации:

- 1 – подпрограмма вычисления эмпирических значений кумулянтов;
- 2 – подпрограмма вычисления значений критерия S ;
- 3 – блок анализа и определения вида закона распределения;
- 4 – блок вычисления (хранения) аналитических значений кумулянтов

3. Пример

Для проверки метода был проведён вычислительный эксперимент. При этом использовался математический пакет Mathcad 14. Суть эксперимента состоит в следующем: с помощью датчиков случайных чисел генерировались реализации случайных чисел, распределённых по гауссовскому, равномерному, экспоненциальному и пуассоновскому законам распределения. При проведении вычислительного эксперимента размер выборки был $l = 2^{17}$. Сначала вычислялись теоретические значения моментов в соответствии с выражениями (4), (9), (12), (18). При этом в качестве параметров законов распределений были выбраны следующие: для распределения Гаусса и равномерного распределения математическое ожидание $m = 0$, дисперсия $D = 1$, для экспоненциального распределения и распределения Пуассона $\lambda = 1$. Затем вычислялись значения эмпирических моментов в соответствии с выражениями (19) и значения эмпирических кумулянтов в соответствии с выражениями (4), (9), (12), (18). Результаты сведены в таблицу 1.

Таблица 1

Значения кумулянтов

Закон распределения		χ_1	χ_2	χ_3	χ_4	χ_5	χ_6
Гауссовский	Теоретические	$\chi_1^z = 0$	$\chi_2^z = 1$	$\chi_3^z = 0$	$\chi_4^z = 0$	$\chi_5^z = 0$	$\chi_6^z = 0$
	Эмпирические	$\chi_1^{ez} = 0,0007$	$\chi_2^{ez} = 1,001$	$\chi_3^{ez} = -0,011$	$\chi_4^{ez} = 0,0008$	$\chi_5^{ez} = -0,0069$	$\chi_6^{ez} = -0,047$
Равномерный	Теоретические	$\chi_1^p = 0$	$\chi_2^p = 1$	$\chi_3^p = 0$	$\chi_4^p = -1,2$	$\chi_5^p = 0$	$\chi_6^p = 6,88$
	Эмпирические	$\chi_1^{pz} = -0,00074$	$\chi_2^{pz} = 1,01$	$\chi_3^{pz} = -0,0021$	$\chi_4^{pz} = -1,203$	$\chi_5^{pz} = -0,013$	$\chi_6^{pz} = 6,862$
Экспоненциальный	Теоретические	$\chi_1^e = 1$	$\chi_2^e = 1$	$\chi_3^e = 2$	$\chi_4^e = 6$	$\chi_5^e = 24$	$\chi_6^e = 120$
	Эмпирические	$\chi_1^{ez} = 0,999$	$\chi_2^{ez} = 0,999$	$\chi_3^{ez} = 2,011$	$\chi_4^{ez} = 6,115$	$\chi_5^{ez} = 25,024$	$\chi_6^{ez} = 121,69$
Пуассона	Теоретические	$\chi_1^n = 1$	$\chi_2^n = 1$	$\chi_3^n = 1$	$\chi_4^n = 1$	$\chi_5^n = 1$	$\chi_6^n = 1$
	Эмпирически	$\chi_1^{nz} = 0,999$	$\chi_2^{nz} = 0,997$	$\chi_3^{nz} = 0,996$	$\chi_4^{nz} = 1,006$	$\chi_5^{nz} = 1,004$	$\chi_6^{nz} = 0,598$

Для каждого набора теоретических и эмпирических кумулянтов по формуле (20) вычислялись значения критерия S_i , которые сведены в таблицу 2.

Таблица 2

Значения критериев S_i для различных законов распределений

Распределение	S_1	S_2	S_3	S_4
Гаусса	$S_{1r}=0,0024$	$S_{2r}=0,55$	$S_{3r}=15028,65$	$S_{4r}=5,13$
Равномерное	$S_{1p}=48,54$	$S_{2p}=0,0006$	$S_{3p}=13433,7$	$S_{4p}=42,24$
Экспоненциальное	$S_{1e}=15477,09$	$S_{2e}=15417,43$	$S_{3e}=3,99$	$S_{4e}=15170,41$
Пуассона	$S_{4n}=4,37$	$S_{4n}=4,9$	$S_{4n}=14811,60$	$S_{4n}=0,16$

Рассмотрим подробно алгоритм расчёта первой строки таблицы 2. Каждая строка таблицы рассчитывалась следующим образом. В первой строке мы анализируем выборку, сгенерированную датчиком случайных чисел, распределённых по гауссовскому закону распределения. Расчёт ведём в соответствии со следующими выражениями, при этом используем результаты расчёта теоретических значений кумулянтов для гауссовского, равномерного, экспоненциального и пуассоновского законов распределения из таблицы 1:

$$S_{1z} = \sum_{j=1}^6 (\chi_j^z - \chi_j^{z*})^2; S_{2z} = \sum_{j=1}^6 (\chi_j^p - \chi_j^{z*})^2; S_{3z} = \sum_{j=1}^6 (\chi_j^e - \chi_j^{z*})^2; S_{4z} = \sum_{j=1}^6 (\chi_j^n - \chi_j^{z*})^2. \quad (21)$$

Аналогичным образом рассчитаны остальные строки таблицы 2. Из приведённой таблицы 2 видно, что для первого эксперимента наименьшее значение критерия S_{1r} соответствует распределению Гаусса. Для второго эксперимента наименьшее значение критерия S_{2p} соответствует равномерному распределению, для третьего – наименьшее значение критерия S_{3e} – экспоненциальному распределению, для четвёртого S_{4n} – распределению Пуассона.

Заключение

Таким образом, новизна предлагаемого метода состоит в том, что впервые для решения задачи разработки информационной технологии, позволяющей полностью автоматизировать процесс идентификации вида закона распределения использован кумулянтный метод анализа результатов измерений.

Проведя вычисления по формулам (4), (9), (12), (15), (18), (19), (20), можно определить вид закона распределения, которым описывается выборка результатов измерений Y_i . Таким образом, разработан новый кумулянтный метод определения вида закона распределения, а на базе его – информационная технология, позволяющая полностью автоматизировать процесс определения вида закона распределения результатов измерений. Правильность гипотезы может быть проверена в дальнейшем с помощью известных критериев, например Колмогорова или Стьюдента.

Список литературы

1. Алешкин А.Н., Лабутин С.А. Идентификация формы закона распределения случайных величин как задача приближения функций // Современные проблемы математики и естествознания. Методы и средства измерений: материалы заочных ВНТК 2002 г. – Н. Новгород: МВВО АТН РФ, 2002. – С. 6–9.

2. Безуглов Д.А., Поморцев П.М., Скляров А.В. Обработка результатов измерений на базе аппроксимации плотности распределения сглаживающими кубическими В-сплайнами // Измерительная техника. – 2000. – № 9. – С. 32–36.
3. Малахов А.Н. Кумулянтный анализ случайных негауссовых процессов и их преобразований. – М.: Сов. радио, 1978. – 376 с.
4. Крамер Г. Математические методы статистики. – М.: ИЛ, 1948. – 648 с.
5. Безуглов Д.А. Кумулянтный метод оценки эффективности сегментированного зеркала адаптивной оптической системы // Оптика атмосферы и океана. – 1996. – № 1. – С. 7.

УДК 681.5:519.87

Н.Б. Клименко, А.И. Трясоруков, А.В. Алабуг

ОБРАБОТКА ИСХОДНЫХ ДАННЫХ ДЛЯ ПРИМЕНЕНИЯ МЕТОДОВ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ В ИСЛЕДОВАНИИ ВЛИЯНИЯ ОСОБЕННОСТЕЙ КОЛЕННОГО СУСТАВА НА НАЛИЧИЕ ПРИЗНАКОВ ЕГО ЗАБОЛЕВАНИЙ

В данной статье приведён подход к решению задач, характеризующихся большим количеством взаимосвязанных параметров, среди значений исходной выборки которых отсутствует система группирования. Приведена последовательность действий, позволяющая выявить коррелирующие параметры и сформировать группы, однородные по значениям. Для таких групп с помощью методов математического моделирования определены аналитические зависимости, достоверно отражающие реальные связи параметров.

Введение. Каждый человек может иметь особенности анатомического строения коленного сустава, которые могут оказывать влияние на наличие признаков заболеваний и степень их тяжести. Выявление наличия таких зависимостей и получение числовых характеристик степени тесноты связи (коэффициентов корреляции) между анатомическими параметрами, а также между ними и признаками заболеваний имеет большое значение при диагностировании и лечении. Определить функциональную зависимость наличия заболеваний от анатомических факторов невозможно в силу сложности и многообразия воздействия всевозможных условий на состояние коленного сустава.

Единственно возможный путь для решения задачи выявления наличия этих зависимостей – это метод статистических исследований. При этом статистический анализ позволяет количественно оценить степень влияния

каждого из факторов на степень болезни, выявить их взаимосвязь и выработать оптимальную методику измерений на рентгенограммах коленного сустава.

Для статистических исследований необходимо собрать большой объём информации.

Исходные данные для статистических исследований были получены на основе рентгенограмм коленного сустава разного типа больных.

Был составлен перечень факторов, которые, предположительно, могут оказывать значимое влияние на наличие, вид и степень заболевания. В каждой рентгенограмме проводились замеры по 46 параметрам.

Отобранные факторы были сгруппированы по форме отражения параметра. Этот критерий был выбран потому, что в зависимости от вида представления фактора были использованы разные методы исследования зависимости. Были определены три группы факторов. Среди них:

- 1 группа – факторы, фиксирующие наличие анатомической характеристики (свойства) или признака;
- 2 группа – факторы, фиксирующие значение некоторого анатомического признака количественно;
- 3 группа – факторы, определяющие наличие признака заболевания, патологии или особенности.

Исходные данные для статистических исследований были получены на 115 рентгенограммах коленного сустава больных разного пола, возраста, телосложения.

Далее надо было определить методики анализа взаимозависимости между анатомическими показателями коленного сустава, полученными по параметрам рентгенограмм, в зависимости от формы их представления. Многие анатомические факторы взаимозависимы между собой. Поэтому наличие такой корреляции надо было исключить при формировании множественных зависимостей, отражающих влияние анатомических факторов на факторы, отражающие признаки заболевания.

При наличии корреляционной зависимости между анатомическими факторами были определены независимые анатомические факторы и аналитический вид регрессионных зависимостей для зависимых анатомических факторов.

Далее определили наличие статистической зависимости между факторами – симптомами и независимыми анатомическими факторами и далее определили аналитический вид таких статистических зависимостей.

В процессе исследования использовались следующие базовые положения.

Методы корреляционного анализа используются для изучения особых связей между явлениями и признаками. Корреляционная связь проявляется в ситуации, когда одному и тому же значению аргумента

(независимой переменной) соответствует ряд значений функции (зависимой переменной). Тогда связь обнаруживается в виде тенденции изменений средних значений функции в зависимости от изменений аргументов.

Характеризующие корреляционную связь признаки делятся на факторные и результативные. Факторными называются признаки, влияющие на результат. Признаки, реагирующие на факторные признаки, называются результативными.

Корреляция может возникнуть между двумя или несколькими признаками. В первом случае она называется парной, во втором – множественной.

Основные задачи корреляционного анализа:

- определение и выражение формы аналитической зависимости результативного признака, назовём его Y , от факторных признаков, одного или нескольких, назовём их $X_1, X_2, \dots, X_k$, где k – число факторных признаков;
- измерение тесноты корреляционной связи.

Математическим средством решения первой задачи корреляционного анализа является метод наименьших квадратов.

Измерение тесноты связи между Y и факторами $X_1, X_2, \dots, X_k$ выполняется с помощью расчёта коэффициента корреляции или коэффициента детерминации.

Полученные корреляционные зависимости используются для решения следующих практических задач.

По имеющимся значениям факторных признаков прогнозируется возможное значение результатного признака. Для этого по исходной совокупности определяется уравнение связи результатного признака и факторных признаков; затем производится выборка единиц новой совокупности, в которой по каждому факторному признаку рассчитываются средние величины, которые подставляются в уравнение связи. Так исчисляется средняя величина результатного признака. При этом надо учитывать, что полученное уравнение связи можно использовать для прогнозирования только в том случае, если диапазоны колебаний значений факторных признаков в новой совокупности не будут существенно отличаться от значений этих признаков в исходной совокупности, по которой определено уравнение связи.

Другая задача является обратной. То есть по заданной величине результатного признака определяют необходимые значения по каждому факторному признаку. Такая задача возникает при отыскании оптимального соотношения между факторными и результативными признаками.

Можно рассматривать корреляцию и в таком виде, когда по всем признакам, кроме одного или нескольких, принимается постоянное значение, например среднее. Тогда рассматривается зависимость результатного признака от недетерминированных факторных признаков в частном виде, т.е. это частная корреляция.

По форме аналитического выражения зависимости различают линейную и нелинейную корреляцию. Эти формы характерны для количественных признаков.

Коэффициент корреляции R_r принимает значение по модулю от 0 до 1 и характеризует степень воздействия на вариацию результатного признака существенных факторов. Коэффициент корреляции R_m также используется для определения точности прогноза значений результатного фактора.

Для практических расчётов тесноты связи между результатным и факторным признаками выполняют расчёт минимальной величины коэффициента корреляции, превышение которой служит основанием для утверждения о существенности полученного коэффициента.

Минимальная величина коэффициента множественной корреляции определяется по формуле:

$$R_{\min} = \sqrt{N - m + 4t^2} - \sqrt{\frac{N - m}{2t}},$$

где N – число наблюдений (объём выборки); m – число параметров уравнения корреляции; $t = R/\sigma_r$, если $t > 2$, то с вероятностью 0,95 можно утверждать значимость коэффициента корреляции.

Для $N=100$ и $t = 2$ $R_{\min}=0,2$ для парной корреляции.

Для $N=100$, $t = 2$, $5 < m < 20$ $0,2 < R_{\min} < 0,21$.

Вышеприведённые методы можно применять в том случае, если исходная совокупность значений признаков является однородной. Однородной принимается обычно выборка, для которой коэффициент вариации (отношение среднеквадратического отклонения к выборочной средней признака) не превышает 33 %.

В противном случае исходную выборку по каждому фактору разбивают на несколько частей. В каждой такой группе собираются значения по наличию однородности одного признака. Для каждой такой группы и определяют регрессионные зависимости.

Как предложено выше, признаком однородности выборки может служить значение коэффициента вариации $v < 33$ %. Эта величина может использоваться при отборе значений признака, после их ранжирования.

Используя вышеприведённые положения, были получены следующие результаты.

В таблице 1, представленной ниже, показаны результаты расчёта коэффициентов парной корреляции между факторами второй группы по 100 рентгенограммам.

Таблица 1

**Результаты расчёта коэффициентов парной корреляции
между факторами второй группы**

	Пок4	Пок5	Пок6	Пок8	Пок9	Пок10	Пок11	Пок12	Пок15	Пок16	Пок17	Пок18	Пок19	Пок20.2	Пок20.3	Пок22.1	Пок22.2	Пок22.3	Пок23	Пок25	Пок26	Пок27	Пок28.1	Пок28.2	Пок29.1	Пок29.2
Пок4	1																									
Пок5	-0,01894	1																								
Пок6	0,72822	-0,3948	1																							
Пок8	-0,03642	0,1121	0,0383	1																						
Пок9	-0,0838	0,231	-0,1342	0,51	1																					
Пок10	-0,02856	-0,1618	0,052	-0,2	-0,098	1																				
Пок11	-0,06749	0,3082	-0,0799	0,62	0,512	-0,036	1																			
Пок12	-0,06331	0,2258	0,0156	0,59	0,682	-0,124	0,773	1																		
Пок15	-0,05554	-0,3193	0,1976	0,06	-0,022	0,247	-0,136	-0,167	1																	
Пок16	0,21316	-0,0796	0,2814	-0,04	-0,091	-0,051	-0,055	0,0014	0,1003	1																
Пок17	0,23782	-0,0833	0,2055	-0,2	-0,109	-0,129	-0,21	-0,193	0,1123	0,774	1															
Пок18	0,00273	0,1449	-0,0021	0,2	0,168	0,024	0,379	0,3151	0,0316	0,047	-0,108	1														
Пок19	0,0528	0,3267	0,0044	0,18	0,238	-0,162	0,262	0,3065	-0,117	-0,02	-0,048	0,2696	1													
Пок20.2	-0,14123	0,1299	-0,1657	0,02	0,09	-0,118	0,312	0,1469	0,0118	0,024	0,059	0,0777	0	1												
Пок20.3	0,13343	-0,1311	0,1591	-0,03	-0,095	0,115	-0,319	-0,156	-0,015	-0,02	-0,055	-0,08	-0	-1	1											
Пок22.1	-0,15763	-0,0988	-0,108	0,05	-0,001	-0,185	-0,039	-0,03	-0,034	-0,15	-0,039	-0,105	-0,05	-0,13	0,128	1										
Пок22.2	-0,16856	-0,0772	-0,0988	0,06	0,007	-0,181	0,095	0,0144	0,0994	-0,11	-0,009	-0,075	-0,02	0,08	-0,08	0,788	1									
Пок22.3	0,00238	0,2357	-0,0727	0,09	0,182	-0,047	0,246	0,2282	-0,053	0,103	0,045	0,4858	0,08	0,23	-0,23	-0,11	-0,18	1								
Пок23	-0,03781	0,2278	-0,0589	0,06	0,274	-0,04	0,221	0,287	-0,139	-0,12	-0,212	0,0498	0,07	-0,01	0,008	-0,13	-0,18	0,048	1							
Пок25	-0,0428	-0,1786	0,1607	-0,15	-0,081	-0,007	-0,136	0,0486	0,066	2E-04	-0,061	-0,139	-0,05	-0,06	0,051	-0,16	-0,13	0,002	0,16	1						
Пок26	0,07549	0,0325	0,0984	0,34	0,187	-0,252	0,331	0,2981	-0,151	0,139	-0,08	0,2649	0,15	0,14	-0,15	0,129	0,043	0,327	0,18	-0,039	1					
Пок27	-0,03015	0,0582	0,0571	0,27	0,154	-0,166	0,263	0,2501	-0,33	0,272	0,073	0,1542	0,23	0,04	-0,04	0,054	-0,07	0,288	0,03	-0,165	0,66	1				
Пок28.1	-0,21568	-0,107	-0,029	0,22	0,019	-0,057	0,086	0,0214	-0,097	-0,17	-0,167	-0,027	-0,1	-0,16	0,151	0,646	0,46	0,021	0,01	-0,226	0,29	0,4359	1			
Пок28.2	-0,22937	0,1209	-0,0698	0,16	0,154	-0,054	0,122	0,1081	-0,208	-0,12	-0,146	0,0071	-0,04	-0,01	0,008	0,231	0,12	0,161	0,14	-0,06	0,36	0,6028	0,693	1		
Пок29.1	0,11555	0,2793	-0,0431	-0,06	0,157	0,066	0,046	0,093	-0,101	0,1	0,064	0,0887	0,14	0,05	-0,05	-0,51	-0,45	0,152	0,18	-0,116	0,02	0,0628	-0,3986	-0,01	1	
Пок29.2	-0,00267	0,1298	-0,017	-0,1	0,208	-0,295	-0,037	0,1202	-0,078	0,119	0,14	-0,084	0,22	0,23	-0,22	-0,24	-0,16	0,152	0,06	0,087	0,09	0,1398	-0,1885	0,08	0,1696	1

Для этого и массив данных приведён в вид, позволяющий оценивать взаимозависимость факторов.

Из таблицы видно, что имеются факторы в группе 2, взаимозависимые между собой и имеющие, следовательно, значимый коэффициент корреляции. Из 26 факторов, по которым выполнялись замеры, выявлены следующие зависимости:

- П6 коррелирует с П4 ($R=0.728$);
- П6 коррелирует с П5 ($R=0.395$);
- П9 коррелирует с П8 ($R=0.51$);
- П11 коррелирует с П5 ($R=0.31$); П8 ($R=0.62$); П9 ($R=0.51$);
- П12 коррелирует с П8 ($R=0.59$); П9 ($R=0.68$);
- П15 коррелирует с П5 ($R=0.32$);
- П17 коррелирует с П16 ($R=0.77$);
- П18 коррелирует с П11 ($R=0.38$); П12 ($R=0.32$);
- П19 коррелирует с П15 ($R=0.33$); П12 ($R=0.31$);
- П20.2 коррелирует с П11 ($R=0.31$);
- П20.3 коррелирует с П11 ($R=0.32$);
- П22.2 коррелирует с П22.1 ($R=0.79$);
- П22.3 коррелирует с П18 ($R=0.49$);

- П26 коррелирует с П8 ($R=0.34$); П11 ($R=0.33$); П22.1 ($R=0.33$);
- П27 коррелирует с П15 ($R=0.33$); П26 ($R=0.66$);
- П28.1 коррелирует с П22.1 ($R=0.64$); П22.2 ($R=0.46$); П27 ($R=0.44$);
- П28.2 коррелирует с П26 ($R=0.36$); П27 ($R=0.61$); П28.1 ($R=0.69$).

На основании анализа зависимых факторов можно утверждать, что среди анатомических факторов 2 группы статистически независимыми являются П4, П5, П8, П9, П16.

Для зависимых факторов был определён вид аналитических регрессионных зависимостей и степень их достоверности.

Вид аналитических формул зависимостей приведён в таблице 2.

Таблица 2

**Аналитические зависимости взаимозависимых
анатомических признаков**

Зависимый признак	Вид формулы	Степень достоверности
П6	$P_6 = 0,6023 \cdot P_4$, $P_6 = 0,8352 \cdot P_4 - 0,5436 \cdot P_5$	0,84 0,89
П11	$P_{11} = 0,736086 \cdot P_8 + 0,2088 \cdot P_9$	0,99
П12	$P_{12} = 0,508 \cdot P_8 + 0,4022 \cdot P_9$	0,99
П26	$P_{26} = 0,1096 \cdot P_8 + 0,028 \cdot P_9$	0,85
П27	$P_{27} = 0,16025 \cdot P_8 + 0,1642 \cdot P_9 -$ $- 0,1886 \cdot P_{15} + 0,2415 \cdot P_{16}$	0,95
П28.1	$P_{28.1} = 0,000824 \cdot P_{26} + 0,10929 \cdot P_{27}$	0,92
П28.2	$P_{28.2} = 0,00205 \cdot P_{26} + 0,24262 \cdot P_{27}$	0,92
П29.1	$P_{29.1} = 1,825735 \cdot P_9$	0,98

После выбора независимых факторов необходимо определить степень однородности выборки их значений. Для этого определили коэффициент вариации V , среднее и дисперсию по выборке каждого независимого фактора. Результаты представлены в таблице 3.

Таблица 3

**Статистические характеристики независимых
количественных факторов**

Идентификатор	Среднее	Среднеквадратическое отклонение	Коэффициент вариации, %	Однородность
П4	6,79	2,54	37,40	
П5	3,23	1,41	43,53	
П8	45,06	4,45	9,87	да
П9	46,48	6,40	13,78	да
П16	14,48	7,56	52,19	

Как видно из таблицы 3, выборки по П8, П9 являются однородными. Выборки по П4, П5, П16 неоднородны и требуют разбиения.

Для формирования однородных групп по П4, П5, П16 используется предложенный ранее критерий. Его суть: ранжируется факторный показатель и отбираются его значения по возрастанию в группу до тех пор, пока коэффициент вариации не превышает 33 %.

По каждому следующему фактору выбор однородных групп выполняется внутри группы, однородной по предыдущему фактору. Так как у нас три фактора имеют неоднородную совокупность исходных данных, то в результате такой группировки получим 8 групп с однородными факторами П4, П5, П16. В каждой из этих групп определим среднее, дисперсию, коэффициент вариации, количество замеров, долю от исходной совокупности. Результаты группировки приведены в таблице 4.

Таблица 4

Однородные группы по признакам П4, П5, П16

Номер группы	Г1	Г2	Г3	Г4	Г5	Г6	Г7	Г8
Диапазон П4	2–9	10–12	2–9	10–12	2–9	10–12	2–9	10–12
Диапазон П5	1–3		4–7		1–3		4–7	
Диапазон П16	5–15				17–35			
Число замеров в группе <i>n</i>	31	7	21	6	19	3	10	3
Доля от общей совокупности	0,31	0,07	0,21	0,06	0,19	0,03	0,1	0,03

Как видно, представительными являются группы: Г1, Г3, Г5, Г7. По ним и будем далее выполнять дальнейший анализ.

Выводы. Предложенное описание исследования показывает, насколько важным этапом является начальная обработка и анализ первичных статистических данных для корректного статистического моделирования и выявления аналитических зависимостей.

Список литературы

1. Мэйндоналд Дж. Вычислительные методы в прикладной статистике: пер. с англ. / под ред. Е.З. Демиденко. – М.: Финансы и статистика, 1988. – 288 с.
2. Сиськов В.Н. Корреляционный анализ в экономических исследованиях. – М.: Статистика, 1975. – 156 с.

УДК 621.396.96

Н.Г. Пархоменко

**ПОЛЯРИЗАЦИОННО-НЕЗАВИСИМОЕ ОЦЕНИВАНИЕ
УГЛОВОГО ПОЛОЖЕНИЯ ИСТОЧНИКОВ
КОГЕРЕНТНЫХ СИГНАЛОВ НА ОСНОВЕ МЕТОДОВ
НЕКВАДРАТИЧНОЙ РЕГУЛЯРИЗАЦИИ**

Для оценивания направлений на источники радиоизлучений когерентных сигналов разработан метод неквадратичной регуляризации, который учитывает векторный характер электромагнитного поля падающих волн и комплексных диаграмм направленности антенных элементов и не зависит от поляризации приходящих волн.

Ключевые слова: поляризация электромагнитного поля, амплитудно-фазовое распределение, неквадратичная регуляризация, разрешение лучей.

Введение. Пространственная локализация источников радиоизлучения с разрешением, превосходящим рэлеевский предел, представляет собой актуальную и интенсивно разрабатываемую задачу. Исторический первый метод высокого разрешения – метод Кейпона [1] (метод максимального правдоподобия) – имеет разрешение, не намного превосходящее рэлеевский предел, но обладает высокой устойчивостью к случайным и систематическим ошибкам. Более совершенные методы, например, методы теплового шума, метод собственных векторов и метод MUSIC, имеют более высокое пространственное разрешение [1–3]. В методах собственных векторов, методе MUSIC и им подобных эксплуатируется идея разделения пространства, в котором определены векторы амплитудно-фазовых распределений на элементах антенной решётки, на сигнальную и шумовую компоненты. При всех своих очевидных достоинствах и логической стройности эти методы плохо приспособлены к разрешению источников, сигналы которых коррелированы. Особенно сильно это проявляется в коротковолновом диапазоне радиоволн. Поле в точке приёма, создаваемое даже одним источником, формируется в результате интерференции нескольких мод (лучей), распространяющихся по различным траекториям. Если сигнал, распространяющийся по различным траекториям, имеет обратную полосу частот, существенно превышающую время задержки между лучами, то принимаемые сигналы оказываются сильно коррелированы и эффективность указанных методов становится проблематичной. Недостаточно полно разработан и вопрос об учёте векторного характера электромагнитного поля, т.е. поляризации приходящих волн.

В настоящее время интенсивно исследуется и развивается метод неквадратичной регуляризации [4] и его модификации [5–10], которые позволяют оценить энергию приходящего сигнала как функцию угла, которая в идеальном случае должна иметь максимумы для коррелированных лучей с близкими направлениями прихода. В этих работах явно или неявно

предполагается, что антенная решётка состоит из антенных элементов, которые согласованы с полем определённой поляризации. Однако это предположение далеко не всегда соответствует действительности. Даже простейшие линейные вибраторные антенны дают отклик на электрическое поле падающей волны, перпендикулярное оси антенны, вследствие дифракции волн на конструктивных элементах антенны и окружающих предметах. В общем случае элементы антенной решётки дают отклик на обе поляризации падающих волн, и для каждой из них имеют свою комплексную диаграмму направленности (ДН). В работе [11] показано, что учёт векторного характера комплексных ДН антенных элементов позволяет повысить точность и достоверность оценок направления прихода одного луча за счёт исключения поляризационных ошибок, которые могут достигать десятков градусов. Кроме того, предложенный в этой работе подход позволяет получить оценку вектора поляризации, который является дополнительным информационным параметром поля падающей волны.

Цель настоящей работы заключается в разработке метода неквадратичной регуляризации, который даёт оценку направлений прихода нескольких лучей с учётом векторного характера комплексных ДН антенных элементов вне зависимости от поляризации приходящих волн.

1. Метод регуляризации

Постановка задачи. Рассмотрим предварительно задачу оценивания направлений прихода в традиционной постановке, не учитывающей векторную природу поля, и преобразуем окончательные формулы для удобства их использования в случае векторного характера комплексных ДН антенных элементов.

Пусть имеется антенная решётка, содержащая N антенных элементов, которая регистрирует поле множества узкополосных сигналов. Источники этих сигналов находятся в дальней зоне антенной решётки. Можно считать поэтому, что фронты приходящих волн плоские. Пусть Θ_m , $m \in [1, M]$ – узлы сетки направлений ожидаемого прихода волн (лучей), в общем случае двумерной, а $\mathbf{J}_0$ – идеальный (теоретический) N -мерный вектор амплитудно-фазового распределения (АФР), т.е. комплексных амплитуд сигналов на антенных элементах. Математическую модель АФР запишем в виде:

$$\mathbf{A}\mathbf{s}_0 = \mathbf{J}_0. \quad (1)$$

Здесь $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_M)$ – фазирующая $N \times M$ матрица, т.е. матрица, составленная из N -мерных векторов $\mathbf{a}_m$, каждый из которых есть отклик антенной решётки на луч с единичной амплитудой, приходящий с направления Θ_m . Конкретный вид векторов $\mathbf{a}_m$ зависит от геометрии антенной

решётки и типа антенных элементов. Вектор $\mathbf{s}_0$, имеющий размерность M , есть распределение комплексных амплитуд приходящих лучей с направлениями прихода Θ_m , создающее АФР $\mathbf{J}_0$.

В реальной ситуации вектор $\mathbf{J}_0$ неизвестен, а известен вектор $\mathbf{J}$, отличающийся от него величиной шумов, складывающихся из принимаемого фона и внутренних шумов приёмных трактов. Подставляя в правую часть (1) вектор измеренного зашумлённого АФР $\mathbf{J}$, приходим к системе уравнений относительно неизвестного и подлежащего оценке вектора комплексных амплитуд источников поля $\mathbf{s}$

$$\mathbf{A}\mathbf{s} = \mathbf{J}. \quad (2)$$

Поскольку $M > N$, система линейных уравнений (2) является недоопределённой и имеет неограниченное множество решений. Следовательно, для однозначной оценки вектора $\mathbf{s}$ требуется наложение дополнительных условий, т.е. регуляризация задачи [12].

Регуляризация заключается в сведении рассматриваемой задачи к поиску экстремума (минимума) функционала вида:

$$G(\mathbf{s}) = G_1(\mathbf{s}) + \gamma G_2(\mathbf{s}). \quad (3)$$

В выражении (3) $G_1(\mathbf{s})$ – евклидова норма [13] вектора невязок системы уравнений (2):

$$G_1(\mathbf{s}) = (\mathbf{s}^+ \mathbf{A}^+ - \mathbf{J}^+) (\mathbf{A}\mathbf{s} - \mathbf{J}) \equiv \|\mathbf{A}\mathbf{s} - \mathbf{J}\|^2,$$

где $(\cdot)^+$ – символ эрмитова сопряжения, γ – параметр регуляризации; $G_2(\mathbf{s})$ – стабилизирующий функционал [12], выбором которого определяется возможность разрешения лучей с угловым расстоянием, меньшим рэлеевского предела.

Вывод основных уравнений. Положим $\xi_m = |s_m|^2$, где s_m – компоненты вектора $\mathbf{s}$. Вектор ξ с компонентами ξ_m представляет собой распределение энергии поля по пространству, положение максимумов которого даёт искомые направления на источники излучения. Будем считать, что стабилизирующий функционал есть функция только от вектора ξ , но не от переменных s_n, s_n^* по отдельности. Дифференцируя (3) по $\mathbf{s}^+$, находим

$$\frac{\partial G(\mathbf{s})}{\partial \mathbf{s}^+} = (\mathbf{A}^+ \mathbf{A} + \gamma \Lambda(\xi)) \mathbf{s} - \mathbf{A}^+ \mathbf{J},$$

где $\Lambda(\xi) = \text{diag}\{\lambda_m(\xi)\}$, $\lambda_m(\xi) = \partial G_2(\xi) / \partial \xi_m$.

Приравняв эту производную нулю, получим систему уравнений относительно $\mathbf{s}$:

$$(\mathbf{A}^+ \mathbf{A} + \gamma \Lambda(\xi)) \mathbf{s} = \mathbf{A}^+ \mathbf{J}. \quad (4)$$

Система уравнений (4), вообще говоря, нелинейна, поскольку её матрица $\mathbf{A}^+ \mathbf{A} + \gamma \mathbf{\Lambda}(\xi)$ в общем случае зависит от искомого вектора $\mathbf{s}$.

Формальное решение системы уравнений (4) записывается в виде:

$$\mathbf{s} = (\mathbf{A}^+ \mathbf{A} + \gamma \mathbf{\Lambda}(\xi))^{-1} \mathbf{A}^+ \mathbf{J}. \quad (5)$$

Упрощение выражения (5) достигается использованием леммы об обращении матриц [14]. Опуская промежуточные выкладки, окончательный результат запишем в виде:

$$\mathbf{s} = \mathbf{\Lambda}(\xi)^{-1} \mathbf{A}^+ (\mathbf{A} \mathbf{\Lambda}(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J}, \quad (6)$$

или в компонентах вектора $\mathbf{s}$

$$s_m = (\lambda_m(\xi))^{-1} \mathbf{a}_m^+ (\mathbf{A} \mathbf{\Lambda}(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J}, \quad m \in [1, M], \quad (7)$$

где $\mathbf{I}$ – единичная матрица размерности $N \times N$.

Случай квадратичной регуляризации. Рассмотрим частный случай регуляризирующего функционала $G_2(\mathbf{s})$, который выбирается как евклидова норма (энергия) вектора $\mathbf{s}$

$$G_2(\xi) = \mathbf{s}^+ \mathbf{s} = \sum_m \xi_m.$$

Тогда $\lambda_m = 1$, матрица $\mathbf{\Lambda}$ оказывается единичной, и выражение (6) приобретает следующий вид:

$$\mathbf{s} = \mathbf{A}^+ (\mathbf{A} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J}.$$

Поскольку матрица $\mathbf{A} \mathbf{A}^+$ невырождена, можно, положив $\gamma = 0$, написать

$$\mathbf{s} = \mathbf{A}^+ (\mathbf{A} \mathbf{A}^+)^{-1} \mathbf{J} = \mathbf{A}^\# \mathbf{J}, \quad (8)$$

где $\mathbf{A}^\#$ – матрица, псевдообратная матрице $\mathbf{A}$ [14]. Таким образом, вектор $\mathbf{s}$ оказывается нормальным псевдорешением системы уравнений (2) с минимальной евклидовой нормой (энергией).

Метод решения основных уравнений. Если зависимость $G_2(\xi)$ от $\mathbf{s}$ не является квадратичной, то λ_m зависит от ξ , и система уравнений (5) становится нелинейной. Для её решения будем использовать итерационный метод. Сформулируем его непосредственно для вектора ξ , который, собственно, и представляет интерес. Взяв квадрат модуля обеих частей уравнения (7), пишем:

$$\xi_m = \left| (\lambda_m(\xi))^{-1} \mathbf{a}_m^+ (\mathbf{A} \mathbf{\Lambda}(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J} \right|^2, \quad m \in [1, M] \quad (9)$$

Тем самым получены уравнения, в которые входит только вектор энергии поля ξ . Теперь, задавшись начальным приближением $\xi^{(0)}$, на итерации с номером $k > 0$ имеем:

$$\xi_m^{(k)} = \left| \left(\lambda_m(\xi^{(k-1)}) \right)^{-1} \mathbf{a}_m^+ \left(\mathbf{A} \Lambda(\xi^{(k-1)})^{-1} \mathbf{A}^+ + \gamma \mathbf{I} \right)^{-1} \mathbf{J} \right|^2.$$

Начальным приближением $\xi^{(0)}$ можно выбрать, например, квадрат модуля нормального псевдорешения (8) или классическую диаграмму направленности антенной решётки:

$$D_m = \left| \mathbf{a}_m^+ \mathbf{J} \right|^2.$$

Естественным условием завершения процесса служит неравенство:

$$\left\| \xi^{(k+1)} - \xi^{(k)} \right\| < \varepsilon,$$

где ε – заданное малое число.

Выбор стабилизирующего функционала. Выбор стабилизирующего функционала $G_2(\xi)$, вообще говоря, неоднозначен и подсказывается содержанием решаемой задачи. Следуя работе [4], положим

$$G_2(\xi) = \sum_m (\xi_m)^{p/2}, \quad (10)$$

где $p \in (0, 2)$.

При $p \geq 1$ $(G_2)^{1/p}$ представляет собой гёльдерову ℓ_p норму вектора $\mathbf{s}$ [13]. Выбор стабилизирующего функционала в виде (10) мотивируется тем, что при $p < 1$ он обеспечивает выделение минимального числа лучей, которое согласуется с экспериментальными данными. Кроме того, дополнительным обоснованием выбора $p < 1$ служит тот факт, что при $p \rightarrow 0$ функционал $G_2(\xi) \rightarrow Q$, где Q – число приходящих лучей с отличной от нуля энергией. Другими словами, минимизируя функционал $G(\mathbf{s})$ (2), мы находим оценки параметров минимального числа лучей, согласующиеся с наблюдаемыми результатами.

Дифференцируя выражение (10) по ξ_m , получим:

$$\lambda_m(\xi) = \partial G_2(\xi) / \partial \xi_m = \frac{p}{2} (\xi_m)^{p/2-1}.$$

Отсюда следует, что если $p < 2$, то $\lambda_m \rightarrow \infty$ при $\xi_m \rightarrow 0$. Эта особенность не влияет на корректность формул (6), (7), поскольку в них входят только величины λ_m^{-1} , которые обращаются в нуль вместе с ξ_m .

2. Метод регуляризации с учётом поляризации

Рассмотрим обобщение задачи оценивания направлений на источники излучения, когда падающие волны имеют произвольную поляризацию, а антенные элементы дают отклик на обе компоненты напряжённости поля и имеют для этих компонент различные комплексные диаграммы направленности. Как и ранее, предположим, что поле нескольких узкополосных сигналов регистрируется антенной решёткой, состоящей из N антенных элементов. Разложим векторы напряжённости поля каждой из падающих волн на две компоненты. Назовём их условно горизонтальной и вертикальной компонентами и будем обозначать верхними индексами h и v . Математическая модель задачи сохраняет форму (1), однако матрица $\mathbf{A}$ приобретает теперь блочную структуру:

$$\mathbf{A} = [\mathbf{A}^{(h)}, \mathbf{A}^{(v)}]$$

и имеет теперь размерность $N \times 2M$. Число узлов сетки направлений Θ_m сохраняется при этом неизменным.

Векторы $\mathbf{a}_n^{(h)}$ – столбцы матрицы $\mathbf{A}^{(h)}$ – и векторы $\mathbf{a}_n^{(v)}$ – столбцы матрицы $\mathbf{A}^{(v)}$ – это отклики антенной решётки на лучи единичной амплитуды соответствующей поляризации, приходящие с направления Θ_m . Компоненты $a_{nm}^{(h)}$ векторов $\mathbf{a}_n^{(h)}$ и компоненты $a_{nm}^{(v)}$ векторов $\mathbf{a}_n^{(v)}$ – это отклики n -й антенны на лучи единичной амплитуды соответствующей поляризации, приходящих с направления Θ_m . Они пропорциональны значениям комплексных диаграмм направленности антенн в направлении Θ_m для горизонтальной и вертикальной поляризаций, соответственно. Конкретный вид элементов $a_{nm}^{(h)}$ матрицы $\mathbf{A}^{(h)}$ и элементов $a_{nm}^{(v)}$ матрицы $\mathbf{A}^{(v)}$ зависит от типа и ориентации антенных элементов, места расположения и геометрии антенной решётки, длины волны и выбора координатной сетки направлений Θ_m . Они могут быть найдены либо по результатам измерений, либо по результатам решения граничной задачи электродинамики с учётом конструктивных особенностей антенной решётки и места её расположения. Вектор неизвестных комплексных амплитуд $\mathbf{s}$ имеет теперь размерность $2M$ и также имеет блочную структуру, обусловленную наличием двух компонент поля:

$$\mathbf{s} = \begin{pmatrix} \mathbf{s}^{(h)} \\ \mathbf{s}^{(v)} \end{pmatrix}.$$

Компоненты $s_m^{(h)}$, $s_m^{(v)}$, $m \in [1, M]$, векторов $\mathbf{s}^{(h)}$ и $\mathbf{s}^{(v)}$ – это комплексные амплитуды двух поляризационных компонент поля лучей, приходящих с соответствующих направлений Θ_m . В отсутствие пространственного шумового фона эти векторы (один или оба) будут иметь ненулевые q -е элементы только в том случае, когда имеются источники в направлениях Θ_q .

Вектор комплексных амплитуд $\mathbf{J}$ сигналов на антенных элементах теперь зависит от комплексных амплитуд двух компонент напряжённости падающих волн.

Задача состоит в том, что необходимо оценить амплитуды приходящих лучей двух рассматриваемых поляризаций как функции их углов прихода – азимута и угла места. Модули этих амплитуд имеют несколько максимумов, соответствующих направлениям прихода лучей, причём если приходящий луч имеет отличные от нуля как горизонтальную, так и вертикальную компоненты поля, то направления максимумов этих двух функций должны совпадать.

Чтобы распространить метод регуляризации на рассматриваемый случай, положим $\xi_m = |s_m^{(h)}|^2 + |s_m^{(v)}|^2$. Стабилизирующий функционал G_2 по-прежнему будем считать функцией вектора ξ . Ненулевыми элементами диагональной $2M \times 2M$ матрицы Λ теперь будут $\lambda_m(\xi) = \lambda_{m+M}(\xi) = \partial G_2(\xi) / \partial \xi_m$. Тогда выражение (7) распадается на две компоненты:

$$\begin{aligned} s_m^{(h)} &= (\lambda_m(\xi))^{-1} (\mathbf{a}_m^{(h)})^+ (\mathbf{A} \Lambda(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J}, \\ s_m^{(v)} &= (\lambda_m(\xi))^{-1} (\mathbf{a}_m^{(v)})^+ (\mathbf{A} \Lambda(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J}. \end{aligned} \quad (11)$$

Складывая квадраты модулей левых и правых частей уравнений (11), получим выражение, обобщающее выражение (9):

$$\xi_m = (\lambda_m(\xi))^{-2} \left\{ \left| (\mathbf{a}_m^{(v)})^+ (\mathbf{A} \Lambda(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J} \right|^2 + \left| (\mathbf{a}_m^{(h)})^+ (\mathbf{A} \Lambda(\xi)^{-1} \mathbf{A}^+ + \gamma \mathbf{I})^{-1} \mathbf{J} \right|^2 \right\}, \quad (12)$$

где $m \in [1, M]$.

Это выражение представляет собой уравнение относительно вектора ξ , которое решается итерационным методом, вполне аналогичным рассмотренному выше. Начальное приближение выбирается при этом в виде:

$$\xi_m^{(0)} = \left| (\mathbf{a}_m^{(h)})^+ (\mathbf{A} \mathbf{A}^+)^{-1} \mathbf{J} \right|^2 + \left| (\mathbf{a}_m^{(v)})^+ (\mathbf{A} \mathbf{A}^+)^{-1} \mathbf{J} \right|^2,$$

которое получается из (12) выбором стабилизирующего функционала вида (10) при $p = 2$.

3. Результаты численного моделирования

Рассмотрим в качестве примера плоскую круговую антенную решётку, состоящую из одиннадцати рамочных антенн с размерами, много меньшими длины волны, расположенных эквидистантно на окружности (рис. 1). При размещении в свободном пространстве диаграммы направленности её элементов могут быть записаны в виде замкнутых аналитических выражений исходя из того, что напряжение на выходе рамочной антенны пропорционально проходящему через рамку магнитному потоку. Соответствующие выражения для элементов матриц $\mathbf{A}^{(h)}$ и $\mathbf{A}^{(v)}$ с точностью до постоянного множителя имеют вид:

$$a_{nm}^{(h)} = \sin \theta_n \cdot \cos(\varphi_n - \alpha_k) \cdot \exp(i \mathbf{k}_m \mathbf{r}_n),$$

$$a_{nm}^{(v)} = (\cos \theta_n \cdot \cos \beta_l + \sin \theta_n \cdot \sin \beta_l \cdot \sin(\varphi_n - \alpha_k)) \cdot \exp(i \mathbf{k}_m \mathbf{r}_n),$$

$$m = k + lK,$$

где θ_n – углы наклона плоскостей антенн, отсчитанные от горизонтальной плоскости; φ_n – азимутальные углы фазовых центров антенн; α_k – узлы сетки направлений по азимуту ($k \in [1, K]$); β_l – узлы сетки направлений по углу места ($l \in [1, L]$); $\mathbf{k}_m$ – волновой вектор луча, приходящего с m -го направления; $\mathbf{r}_n$ – радиус-вектор фазового центра n -й антенны.

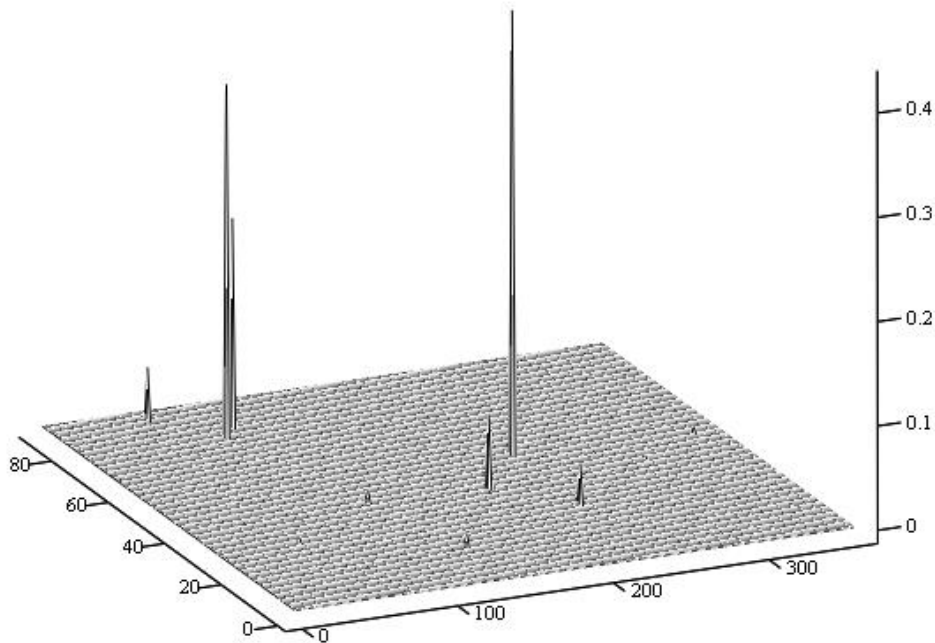


Рис. 1. Результат пеленгования двух лучей с эллиптической поляризацией и азимутальным разносом, превосходящим ширину области разрешения

Рэлея, методом регуляризации без учёта векторного характера поля падающих волн. Азимутальные и угломестные координаты двух лучей с наибольшими амплитудами: (229, 49) и (89, 73) соответственно

На рисунках 1–4 представлены результаты моделирования антенной решётки радиусом 2λ ; λ – длина волны. При моделировании полагалось, что плоскости антенн наклонены к центру окружности под одинаковыми углами так, что $\theta_n = 60^\circ$. Такой наклон улучшает свойства антенной решётки при пеленговании лучей с малым углом места по сравнению с вертикальным расположением плоскостей рамок. Число падающих лучей принято равным двум. Моделирование проведено по одному временному отсчёту (АФР) полностью когерентных сигналов при высоком отношении сигнал/шум. Шаг координатной сетки составлял 1° как по азимуту, так и по углу места.

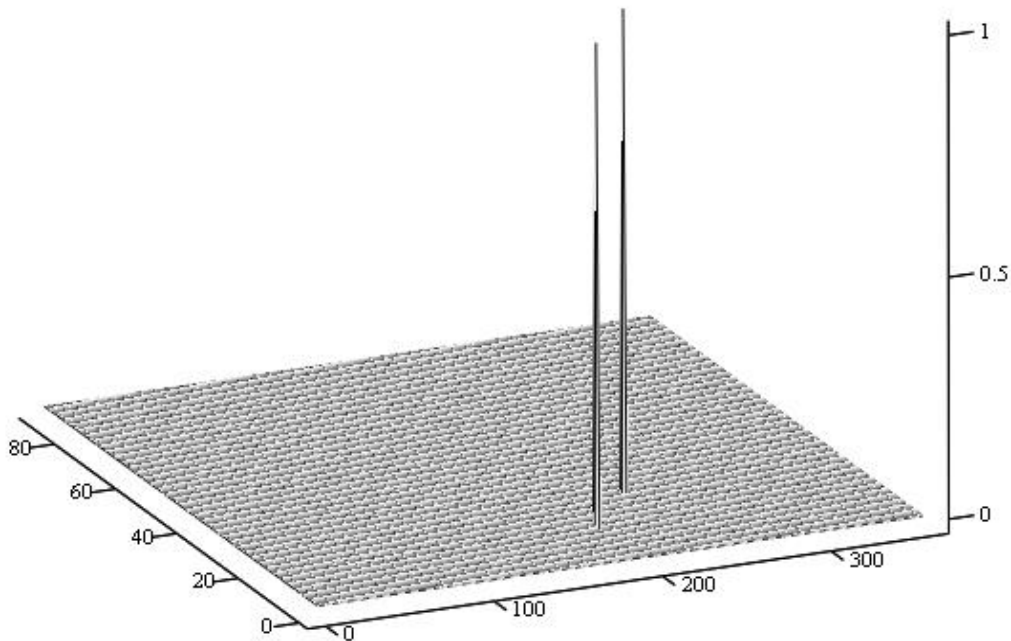


Рис. 2. Результат пеленгования двух лучей с эллиптической поляризацией и азимутальным разносом, превосходящим ширину области разрешения Рэлея, методом регуляризации с учётом векторного характера поля падающих волн. Азимутальные и угломестные координаты двух лучей с наибольшими амплитудами: (227, 25) и (191, 14) соответственно

На рисунках 1 и 2 заданный азимут первого луча составляет 191° , угол места – 14° , комплексная амплитуда горизонтальной компоненты магнитного поля $\exp(i \cdot \pi/4) \cdot \cos(\pi/3)$, вертикальной компоненты – $\exp[i \cdot (\pi/4 + \pi/9)] \cdot \sin(\pi/3)$ (эллиптическая поляризация). Для второго луча азимут составляет 227° , угол места – 25° , комплексные амплитуды горизонтальной и вертикальной компонент равны $1/\sqrt{2}$, сдвиг фазы между ними – $\pi/4$ (эллиптическая поляризация).

Разнос лучей по азимуту превосходит ширину области разрешения Рэлея, которая для данной решётки составляет примерно 30° .

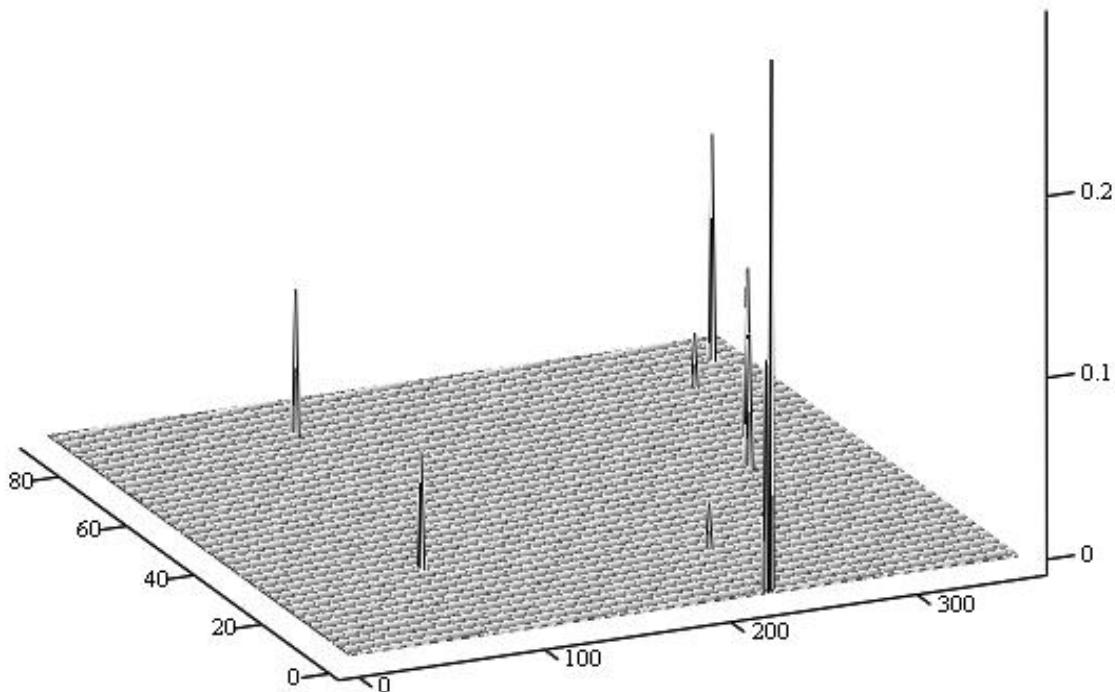


Рис. 3. Результат пеленгования двух лучей с эллиптической поляризацией и азимутальным разносом, меньшим ширины области разрешения Рэлея, методом регуляризации без учёта векторного характера поля падающих волн. Азимутальные и угломестные координаты двух лучей с наибольшими амплитудами: (227, 0) и (224, 0) соответственно

На рисунках 3 и 4 азимут первого луча составляет 225° , угол места – 14° . Для второго луча азимут 227° , угол места – 25° . Комплексные амплитуды горизонтальных и вертикальных компонент магнитного поля лучей такие же, как и для рисунков 1 и 2. Разнос лучей по азимуту существенно меньше ширины области разрешения Рэлея.

На рисунках 1 и 3 приведены результаты расчёта пеленгов, полученные в предположении о том, что приходящие волны имеют строго вертикальную поляризацию (имеют только горизонтальную компоненту магнитного поля). Для проведения расчётов использовался метод регуляризации без учёта векторного характера (поляризации) электромагнитного поля и комплексных ДН антенных элементов, т.е. использовалась формула (9), в которой полагалось $\mathbf{A} = \mathbf{A}^{(h)}$.

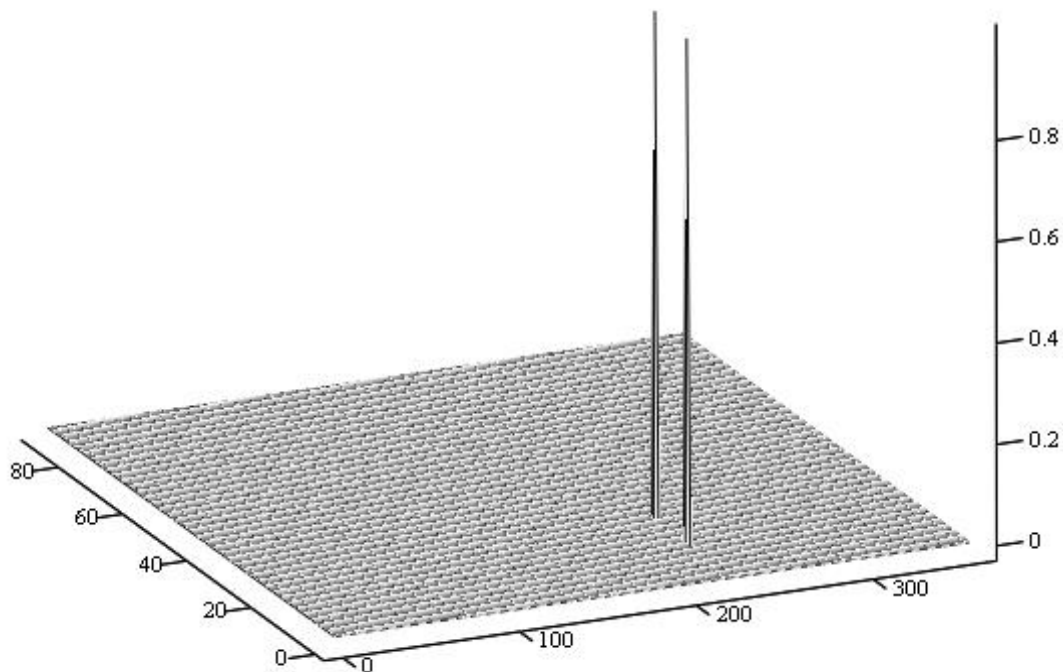


Рис. 4. Результат пеленгования двух лучей с эллиптической поляризацией и азимутальным разносом, меньшим ширины области разрешения Рэлея, методом регуляризации с учётом векторного характера поля падающих волн. Азимутальные и угломестные координаты двух лучей с наибольшими амплитудами: (227, 25) и (225, 14) соответственно

На рисунках 2 и 4 даны результаты, полученные с учётом векторного характера электромагнитного поля падающих волн и комплексных ДН антенных элементов с помощью описанного выше метода (с помощью итераций по формуле (12)).

В подрисуночных подписях рисунках 1–4 перечислены азимуты и углы места двух наибольших по амплитуде максимумов пеленгационных диаграмм. Все расчёты проведены при $p = 0,1$. При $p = 1$ результаты, получаемые даже с учётом векторного характера полей, аналогичны приведённым на рисунках 1 и 3, т.е. не приводят к решению задачи разделения лучей при любом разносе между ними.

Заключение

Рассмотрена задача о локализации нескольких источников радиоизлучения, находящихся в дальней зоне антенной решётки, элементы которой реагируют на поле падающих волн двух поляризаций. Метод регуляризации, который ранее использовался в случае антенных элементов, реагирующих только на поле одной поляризации, модифицирован так, что позволяет учесть векторный характер полей падающих волн и диаграмм

направленности антенных элементов. Возможности предложенной модификации продемонстрированы на примере антенной решётки, образованной рамочными антеннами, расположенными эквидистантно на окружности, причём плоскости рамок наклонены к центру окружности. Полученные результаты подтверждают возможность использования предложенной модификации метода регуляризации для антенных элементов, которые реагируют на обе поляризации сигнала. При этом предпочтительным оказывается использование регуляризирующего функционала, построенного для $p < 1$.

Список литературы

1. Джонсон Д.Х. Применение методов спектрального оценивания к задачам определения угловых координат источников излучения // ТИИЭР. – 1982. – Т. 70, № 9. – С. 126–139.
2. Schmidt R.O. Multiple emitter location and signal parameter estimation // IEEE Trans. – 1986. – AP-34, № 3. – Pp. 276–280.
3. Ратынский М.В. Адаптация и сверхразрешение в антенных решётках. – М.: Радио и связь, 2003. – 200 с.
4. Cetin M., Malioutov D.M., Willsky A.S. A variational technique for source localization based on a sparse signal reconstruction perspective. – In Proc. of the Internat. Conf. on Acoust. Speech and Sign. Processing, Orlando, FL, May 2002. – Vol. 3. – Pp. 2965–2968.
5. Шевченко В.Н. Разделение многочастотного многолучевого поля методами регуляризации // Изв. вузов. Радиофизика. – 2003. – Т. 46, № 2. – С. 150–161.
6. Шевченко В.Н. Вариационные методы оценивания параметров многолучевого поля. Автометрия. – 2003. – Т. 41, № 4. – С. 24–30.
7. Шевченко В.Н. Оценивание углового положения источников когерентных сигналов на основе методов регуляризации // Радиотехника. – 2003. – № 9. – С. 3–10.
8. Иванов Н.М., Рейзенкинд Я.А. Оценивание направлений на источники радиоизлучения методом нелинейного псевдообращения // RLNC*2004: X Междунар. конф., 13–15 апр., 2004, Воронеж – Т. 1. – С. 40–45.
9. Шевченко В.Н., Иванов Н.М., Звездина Ю.А. Повышение эффективности определения направления прихода коррелированных сигналов методом неквадратичной регуляризации // Электромагнитные волны и электронные системы. – 2006. – № 5. – С. 47–50.
10. Шевченко В.Н., Иванов Н.М., Звездина Ю.А. Метод неквадратичной регуляризации пространственных спектров с фильтрацией ложных составляющих // Автометрия. – 2007. – Т. 43, № 1. – С. 5–9.

11. Иванов Н.М., Онищенко В.С., Шевченко В.Н. Пространственная локализация источников поляризованного электромагнитного поля // Радиотехника и электроника. – 2010. – Т. 55, № 1. – С. 49–56.
12. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. – М.: Наука. ГРФМЛ, 1974. – 223 с.
13. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. – М.: Наука. ГРФМЛ, 1968. – 496 с.
14. Марпл.-мл. С.Л. Цифровой спектральный анализ и его применение: пер. с англ. – М.: Мир, 1990. – 584 с.

УДК 621.396.96

Н.Г. Пархоменко

ПОТЕНЦИАЛЬНЫЕ ХАРАКТЕРИСТИКИ РАДИОЛОКАЦИОННЫХ СИСТЕМ С ПОСТОРОННИМ «ПОДСВЕТОМ»

Рассмотрены потенциальные характеристики пассивных радиолокационных систем, использующих для подсвета целей электромагнитные поля, создаваемые спутниковыми радионавигационными и радиовещательными передатчиками, а также наземными передатчиками теле- и радиовещания. Сформулированы общие требования к источникам сигналов подсвета. Показаны преимущества передатчиков цифрового теле- и радиовещания наземного базирования.

Ключевые слова: пассивная радиолокация с посторонним «подсветом», сигнал подсвета, функция неопределённости, характеристики обнаружения и разрешения.

Введение. В последние годы получает практическое развитие технология радиолокации с посторонним подсветом от телевизионных (ТВ) и радиовещательных передатчиков – пассивная радиолокация с посторонним «подсветом» или полуактивная радиолокация. Основные преимущества полуактивных радиолокационных систем (ПАРЛС) обусловлены отсутствием собственного радиопередатчика. При этом достигается абсолютная радиотехническая скрытность, не требуется выделения дефицитного радиочастотного ресурса, отсутствует вредное воздействие на экологию окружающей среды и человека. Эти РЛС не создают взаимных помех, более экономичны в производстве и эксплуатации по сравнению с активными системами. Большие высоты поднятия антенн и мощности используемых широковещательных передатчиков создают благоприятные условия для обнаружения маловысотных и малоразмерных целей. Кроме того, двухпозиционная топология ПАРЛС помогает обнаруживать летательные аппараты, изготовленные по технологии «стэлс», поскольку бистатическая ЭПР этих аппаратов, как правило, выше, чем при обратном рассеянии.

ПАРЛС в общем случае представляет собой многопозиционную систему, состоящую из одного или нескольких источников сигнала подсвета и одной или нескольких приёмных станций, разнесённых в пространстве. Последние связаны между собой в единую систему обработки радиолокационной информации.

Элементарной структурной единицей такой РЛС является бистатическая система (рис. 1). Обнаружение сигнала, отражённого от цели, принятого основной (как правило, направленной) антенной, осуществляется корреляционным методом, причём опорный сигнал для коррелятора принимается вспомогательной антенной напрямую от источника подсвета. При обнаружении отражённого сигнала фиксируется азимут цели $\Theta_{\text{ц}}$ по угловому положению основной антенны, определяется время запаздывания t_3 , отражённого от цели сигнала относительно опорного сигнала, и доплеровский сдвиг частоты отражённого сигнала $\Omega_{\text{дс}}$.

Координаты цели определяются с помощью угломерно-эллиптического метода [4]. При этом по времени запаздывания вычисляется суммарная дальность:

$$r_{\Sigma\text{ц}} = r_1 + r_{\text{ц}} = ct_3 + d, \quad (1)$$

где r_1 – расстояние от источника подсвета до цели; $r_{\text{ц}}$ – расстояние от цели до приёмной антенны РЛС; c – скорость света; d – расстояние от источника подсвета до РЛС (база).

Расстояние до цели определяется выражением:

$$r_{\text{ц}} = \frac{r_{\Sigma\text{ц}}^2 - d^2}{2(r_{\Sigma\text{ц}} + d \cos \Theta_{\text{ц}})}. \quad (2)$$

По доплеровскому сдвигу частоты отражённого сигнала вычисляется радиальная скорость цели (проекция вектора скорости на бистатическую ось, показанную штрихом на рис. 1):

$$V_{\ddot{o}.\ddot{o}\ddot{a}\ddot{a}} = V_{\ddot{o}} \cos \varphi = \frac{\lambda}{4\pi} \frac{\Omega_{\ddot{a}\ddot{a}}}{\cos(\beta/2)}, \quad (3)$$

где λ – длина волны сигнала подсвета; $V_{\text{ц}}$ – модуль вектора скорости цели; φ – угол между вектором скорости цели и биссектрисой бистатического угла; β – величина бистатического угла.

К основным проблемам ПАРЛС относится, во-первых, выбор источников сигнала подсвета, обеспечивающих приемлемые радиолокационные характеристики, во-вторых, подавление прямого сигнала (рис. 1), превышающего полезный сигнал по уровню на 60–90 дБ [5].

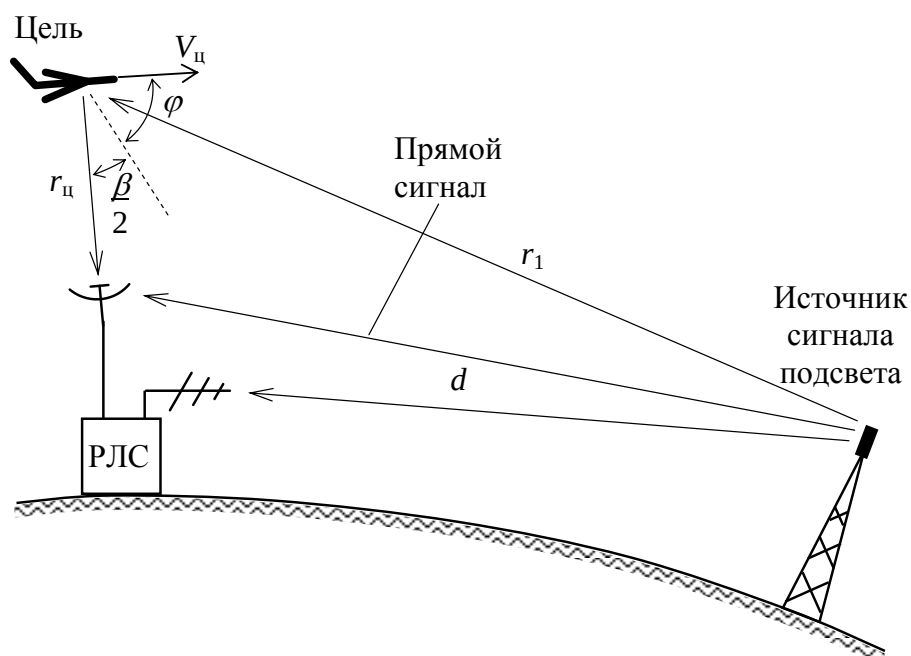


Рис. 1. Бистатическая РЛС с посторонним подсветом

1. Требования к источникам сигнала подсвета

В настоящее время на территории любого государства существует огромное количество источников электромагнитных волн. На населённых территориях эксплуатируется практически весь имеющийся электромагнитный ресурс. Далеко не все источники могут использоваться для радиолокационного подсвета. Требования к источникам подсвета ПАРЛС можно разделить на следующие категории.

Пространственное покрытие определяется формой ДН антенны и излучаемой мощностью. Для радиолокационного подсвета наиболее подходят источники, создающие сплошное радиолокационное поле на большой территории, а именно наземные передатчики со слабонаправленной антенной в азимутальной плоскости, а также авиационные и спутниковые системы. Мощность таких передатчиков должна быть достаточной для обнаружения вторичного излучения целей. Предполагаемые плотности потока мощности излучения у цели [6], рассчитанные для условий распространения в свободном пространстве, показаны на рисунке 2.

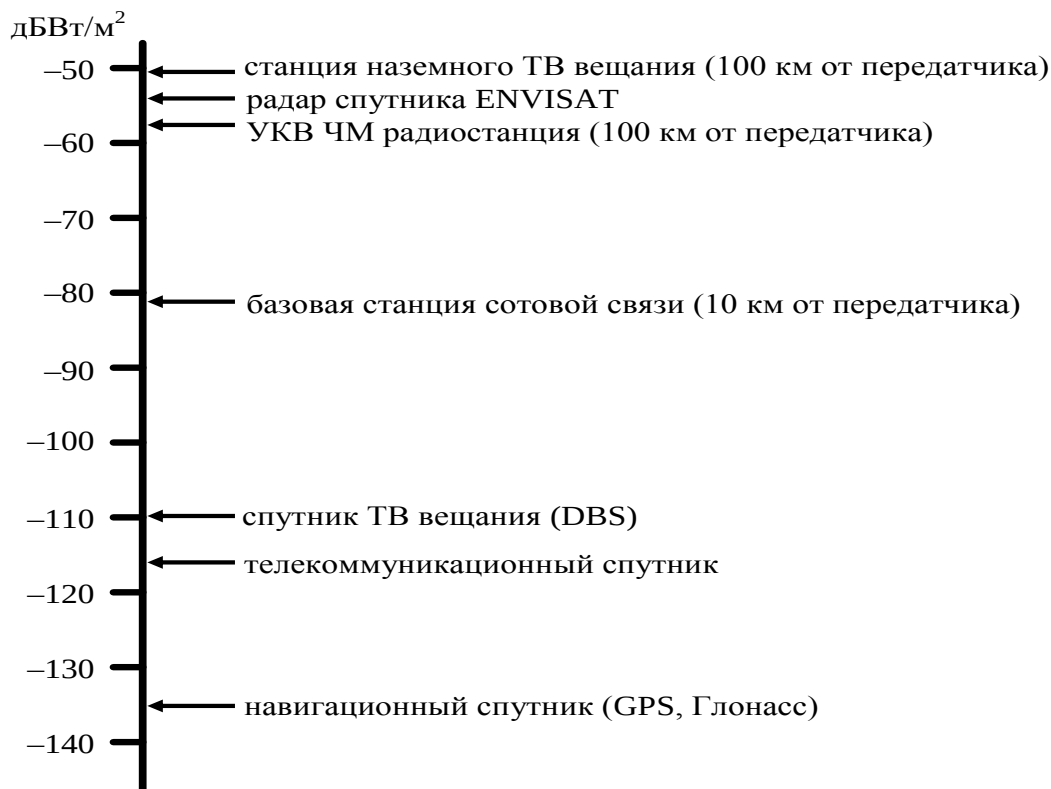


Рис. 2. Плотности потока мощности, создаваемые различными источниками подсвета

Временное покрытие определяется графиком работы передатчика, используемого для подсвета, на территории, обслуживаемой ПАРЛС. С этой точки зрения выгодно использовать системы с круглосуточным вещанием. Для авиационных источников и источников, расположенных в космическом пространстве на околоземной орбите, необходимо учитывать перемещение области подсвета на поверхности Земли. Например, достаточно мощное излучение, создаваемое РЛС разведывательного спутника ENVISAT, малоприспособно для скрытной радиолокации, поскольку спутник осуществляет последовательный обзор всей поверхности Земли и каждый её участок попадает в зону подсвета не чаще одного раза в несколько дней [7].

Частотный диапазон определяется классом контролируемых целей. В общем случае частоты сигналов подсветки могут охватывать рэлеевские и резонансные области вторичного излучения целей. В рэлеевских областях проявляется дифракционное огибание целей волнами подсветки и принимаемый сигнал существенно ослабляется. В резонансных же случаях отражённые сигналы возрастают. Искусственные меры имитации и маскировки целей, эффективные, например, в сантиметровом диапазоне волн, на дециметровых и метровых волнах могут оказаться неэффективными [8].

С точки зрения реализации аппаратуры ПАРЛС, предпочтительным является использование излучений метрового и дециметрового диапазонов. На более низких частотах возрастают размеры антенных систем. На более высоких частотах увеличиваются потери при распространении радиоволн, ухудшаются шумовые характеристики приёмного тракта.

Функция неопределённости (ФН) сигнала подсвета определяет такие характеристики РЛС, как разрешение по дальности и радиальной скорости цели, однозначность определения данных параметров. Боковые лепестки ФН влияют на характеристики обнаружения сигнала. Идеальной является ФН кнопочного типа, характерная для непрерывных шумоподобных сигналов.

Стабильность характеристик сигнала определяет постоянство во времени характеристик обнаружения и разрешения ПАРЛС. Наилучшей стабильностью обладают цифровые системы, в которых модуляция несущей осуществляется псевдослучайными сигналами или рандомизированными цифровыми информационными потоками. Примером источника крайне нестабильного сигнала подсвета является ЧМ радиовещательная станция. Ширина спектра этого сигнала может изменяться в течение трансляции в 2–3 раза при передаче музыки и в 10 и более раз при передаче речи (в паузах между словами) [6]. Для ПАРЛС это означает такие же колебания разрешающей способности по дальности, а также мощности отражённого сигнала (если спектральную плотность принять постоянной).

Рассмотрим возможные источники радиолокационного подсвета.

2. Характеристики возможных источников радиолокационного подсвета

Аналоговое телевизионное вещание. Сигналы наземного аналогового телевизионного (ТВ) вещания рассматривались как наиболее пригодные для радиолокационного подсвета со времён создания первых ПАРЛС и до появления служб цифрового радиовещания.

В спектре полного ТВ сигнала (рис. 3) можно выделить три наиболее энергоёмкие составляющие: несущая сигнала яркости, модуляционная составляющая сигнала яркости, включая сигнал цветности, и сигнал звукового сопровождения (ЗС).

Несущая является удобным средством обнаружения отражённого от цели сигнала при неизвестном времени запаздывания t_z и однозначного определения доплеровского сдвига частоты отражённого сигнала с разрешением, обратно пропорциональным времени его когерентного накопления. Определение времени запаздывания сигнала невозможно.

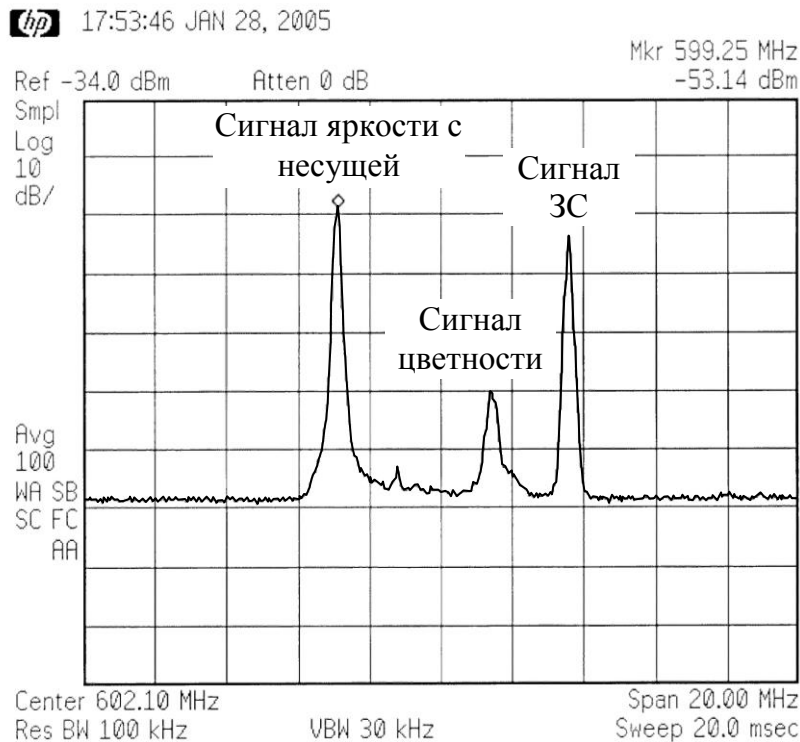


Рис. 3. Спектр полного ТВ сигнала, принятый с расстояния 13 км на комнатную антенну и измеренный спектроанализатором Hewlett-Packard

Амплитудно-модулированный сигнал яркости является непрерывным периодически коррелированным сигналом, и его ФН имеет многолепестковый по задержке характер (рис. 4). Сигнал яркости может использоваться для обнаружения, неоднозначного определения времени запаздывания отражённого сигнала (дальности цели) и однозначного определения доплеровского сдвига частоты отражённого сигнала (радиальной скорости цели).

Анализ ФН (табл. 1) и спектров ТВ сигналов яркости показывает, что его параметры в среднем изменяются в течение трансляции в следующих пределах: время корреляции внутрискановых флуктуаций $\tau_{0я} = 2 \div 5$ мкс, эффективная ширина спектра $\Delta f_{0я} = 100 \div 250$ кГц. Следовательно, потенциальная разрешающая способность по дальности $\Delta r = c\tau_{0я} = c/2\Delta f_{0я} = 600 \div 1500$ м. В связи с высокой корреляцией строк ТВ сигнала, период повторения которых равен $T_{\Pi} = 64$ мкс, диапазон однозначного определения дальности ограничен величиной $r_{одн} = cT_{\Pi}/2 = 9,6$ км, что является существенным недостатком данного

сигнала подсвета. Для устранения этого недостатка могут быть использованы методы фильтрации первичных оценок координат цели. Также предложен метод декорреляции ТВ сигнала яркости путём череспериодного вычитания, который эффективно устраняет неоднозначность, но приводит к существенным энергетическим потерям сигнала [9].

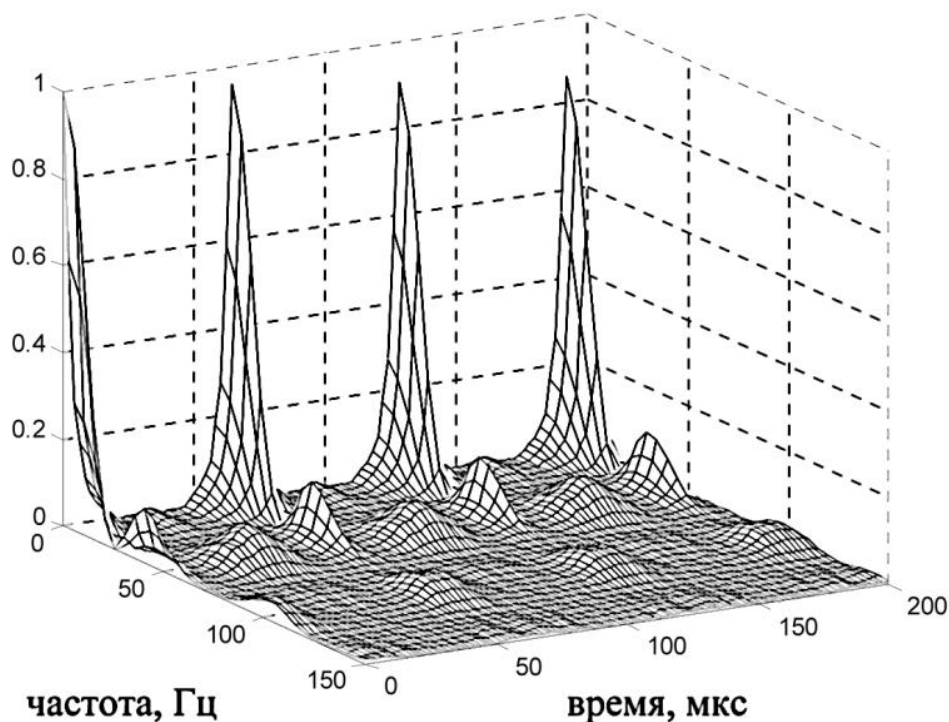


Рис. 4. Типичный вид функции неопределённости ТВ сигнала яркости

Таблица 1

Параметры функций неопределённости ТВ сигнала яркости, рассчитанные по результатам обработки 27 изображений

Параметр	Мин.	Средн.	Макс.	СКО
Ширина главного лепестка по τ , мкс	4	7,69	17	3,47
Максимальный уровень боковых лепестков, дБ	-13	-8,27	-4	2,79
Средний уровень боковых лепестков, дБ	-22,33	-17,08	-10,81	2,99
Коэффициент корреляции строк ТВ СЯ	0,837	0,967	0,995	0,033

Частотно модулированный сигнал звукового сопровождения является непрерывным непериодическим сигналом. ФН этого сигнала, как правило, имеет один выраженный максимум (рис. 5), поэтому возможно однозначное определение времени запаздывания и доплеровского сдвига частоты отражённого сигнала.

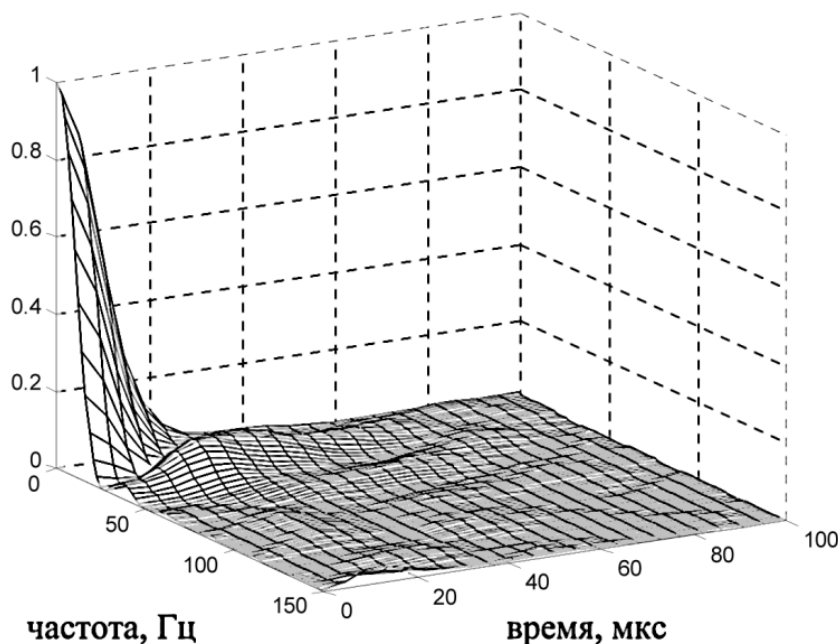


Рис. 5. Типичный вид функции неопределённости ЧМ сигнала звукового вещания

Измерения показывают, что эффективная ширина спектра сигнала звукового сопровождения может изменяться в пределах $\Delta f_{0\text{зс}} = 10 \div 100$ кГц, что подтверждается и результатами анализа ФН (табл. 2). Разрешение по дальности на порядок хуже, чем у сигнала яркости, и равно $\Delta r = c/2\Delta f_{0\text{зс}} = 1,5 \div 15$ км.

ФН сигналов яркости и звукового сопровождения имеют достаточно большой уровень боковых лепестков, достигающий в среднем $C_{\text{бок}} = -8$ дБ, что приводит к ухудшению характеристик обнаружения, как видно из формулы для расчёта отношения сигнал/помеха без учёта мешающих отражений [10]:

$$\rho = \frac{PT_{\text{н}}(1 - C_{\text{бок}}^2)}{N_0 + PT_{\text{н}}C_{\text{бок}}^2 + \frac{P_{\text{пр}}T_{\text{н}}\rho(t_{\text{з}}, F_{\text{дс}})}{\nu}}, \quad (4)$$

где P и $P_{\text{пр}}$ – мощность принятого и прямого сигналов, соответственно; $T_{\text{н}}$ – время когерентного накопления сигнала; $C_{\text{бок}}$ – уровень боковых лепестков корреляционной функции сигнала; $\rho(t_{\text{з}}, F_{\text{дс}})$ – уровень ФН при заданном времени запаздывания и доплеровском сдвиге частоты принятого сигнала; N_0 – спектральная плотность шума; ν – эффективность подавления прямого сигнала специальными средствами РЛС.

Таблица 2

**Параметры функций неопределённости ЧМ
сигнала звукового вещания, рассчитанные по результатам обработки
21 звукового фрагмента**

Параметр	Мин.	Средн.	Макс.	СКО
Ширина главного лепестка по τ , мкс	12	19,8	48	8,5
Максимальный уровень боковых лепестков, дБ	-13	-10,5	-2	2,9
Средний уровень боковых лепестков, дБ	-22,85	-20,09	-14,93	2,05

При использовании аналогового ТВ вещания как источника сигнала подсвета ПАРЛС наилучшие характеристики обнаружения и разрешения могут быть достигнуты при комплексировании обработки всех трёх составляющих.

Радиовещание в диапазоне КВ. Отдельные результаты экспериментальных исследований по изучению спектров ионосферных сигналов вещательных станций в окрестности несущей частоты, а также по обнаружению самолётов с их использованием в качестве подсветки представлены в работах [11, 12]. Показано, что резонансные частоты широкого класса самолётов (пассажирских, транспортных, стратегических бомбардировщиков, штурмовиков, истребителей) лежат в диапазоне от 3 до 20 МГц. Для практического применения ионосферных сигналов в качестве подсветки в ПАРЛС требуются дальнейшие исследования.

Радиовещание с частотной модуляцией в диапазоне УКВ. Пространственное покрытие радиовещательных станций с частотной модуляцией (ЧМ) в диапазоне УКВ охватывает практически все населённые территории, что позволяет организовать радиолокационное наблюдение в любом месте. Вещание осуществляется, как правило, круглосуточно. Использование данного сигнала в интересах радиолокации было осуществлено в ПАРЛС «Silent Sentry», разработанной корпорацией Lockheed Martin в США [13].

ФН аналогична сигналу звукового ТВ сопровождения (см. рис. 5). Радиосигнал УКВ ЧМ вещания может использоваться для однозначного определения времени запаздывания и доплеровского сдвига частоты отражённого сигнала. Из-за нестабильности параметров сигнала разрешение по дальности цели может изменяться в пределах 1,5÷15 км, причём для новостных радиопрограмм оно в среднем составляет 15 км, а для музыкальных – 1,5÷6 км [6].

Для УКВ ЧМ радиовещания выделены диапазоны частот 66÷74 МГц и 87,5÷108 МГц. Преимущество сигнала звукового ТВ сопровождения состоит в том, что он транслируется на более высоких частотах.

Вещательные и исследовательские спутники. С точки зрения энергетических и спектрально-временных характеристик, наибольший интерес для наземных ПАРЛС могли бы представлять радиоизлучения,

создаваемые радиолокационной аппаратурой исследовательских спутников, расположенных на околоземной орбите, однако они имеют, как правило, большой период обращения и не смогут обеспечить непрерывный подсвет целей ПАРЛС. Например, период обзора Земного шара для спутника ENVISAT составляет 1–3 дня.

Вещательные спутники, размещённые на геостационарной орбите, могут обеспечить непрерывное радиолокационное поле у поверхности Земли плотностью порядка $\Pi = -110$ дБВт/м² (см. рис. 2). Отношение сигнал/шум на входе обнаружителя ПА РЛС для свободного пространства оценивается выражением:

$$\rho = \frac{\Pi \sigma_{\text{ц}} S_{\text{а}} T_{\text{н}}}{4\pi r_{\text{ц}}^2 N_0}, \quad (5)$$

где $\sigma_{\text{ц}}$ – эффективная отражающая поверхность цели; $S_{\text{а}}$ – эффективная площадь раскрыва антенны РЛС; $T_{\text{н}}$ – время когерентного накопления; $r_{\text{ц}}$ – дальность цели; N_0 – спектральная плотность шума.

Расчёт по формуле (5) показывает, что при $\sigma_{\text{ц}} = 1$ м², $S_{\text{а}} = 1$ м², $T_{\text{н}} = 100$ мс, $N_0 = 2 \cdot 10^{-20}$ Вт/Гц отношение сигнал/шум $\rho = 50$ будет достигаться только для целей с дальностью $r_{\text{ц}} < 270$ м.

Таким образом, мощность радиовещательных, а тем более радионавигационных спутников недостаточна для подсвета ПАРЛС. Экспериментальные результаты обнаружения надводных и наземных объектов с использованием сигналов навигационных систем ГЛОНАСС и GPS приведены в работе [14].

Службы цифрового радиовещания. Эфирное вещание радио и ТВ программ во всех частотных диапазонах неизбежно будет переведено в цифровой формат в силу ряда объективных причин, важнейшей из которых является высокая спектральная эффективность цифровых методов передачи. Наземное эфирное радиовещание в цифровом формате находится в стадии развёртывания. Для ПА радиолокации наибольший интерес представляет стандарт DAB, предназначенный для организации цифрового звукового радиовещания в диапазонах от 30 МГц до 3 ГГц, и стандарт DVB-T цифрового ТВ вещания.

Радиосигнал, согласно европейскому стандарту DAB [15], формируется путём квадратурной модуляции множества ортогональных несущих колебаний рандомизированными цифровыми потоками данных (OFDM-модуляция). В результате образуется шумоподобный сигнал со спектром, по форме близким к прямоугольному, шириной $\Delta f_{\text{0DAB}} = 1,536$ МГц. Разрешающая способность по дальности составит $\Delta r = c/2\Delta f_{\text{0DAB}} = 98$ м.

Сигнал цифрового ТВ вещания по стандарту DVB-T [16] также формируется на основе OFDM модуляции и имеет ширину спектра $\Delta f_{\text{оц}} = 7,61$ МГц, что обеспечит потенциальную разрешающую способность целей по дальности $\Delta r = c/2\Delta f_{\text{оц}} = 19,6$ м.

С точки зрения полуактивной радиолокации, цифровое телевидение обладает наибольшими перспективами как по радиолокационным характеристикам, так и по масштабам охвата территорий и срокам ввода в эксплуатацию. В настоящее время для цифрового ТВ вещания выпускаются передатчики с выходной мощностью от единиц ватт до 20 кВт. При этом основа вещательной сети строится с использованием мощных передатчиков (1 кВт и более), а маломощные устройства используются для расширения зоны покрытия, улучшения приёма в теневых зонах. Все передатчики предназначены для работы в рамках существующих ТВ каналов метрового и дециметрового диапазонов.

Сигнал цифрового ТВ может формироваться в режиме «2К» с 1705 ортогональными несущими или в режиме «8К» с 6817 несущими. Режим «8К» является основным. Большая часть несущих модулируется данными, которые можно считать случайным некоррелированным потоком. Некоторая часть несущих отводится под передачу непрерывных и распределённых пилот-сигналов, мощность которых повышена на 2,5 дБ, а также служебной информации (TPS). Отдельно формируемые фрагменты сигнала называются символами. Каждый символ содержит полезный и защитный интервалы. В последнем дублируется часть сигнала полезного интервала.

Наличие пилот-сигналов и защитного интервала обуславливает появление в корреляционной функции опасных по уровню составляющих (рис. 6), наибольшая из которых обусловлена защитным интервалом и имеет уровень от -14 дБ при относительной длительности защитного интервала 1/4 до -30 дБ при относительной длительности 1/32. Режекция мешающих составляющих, предложенная в [18], может улучшить характеристики обнаружения при сравнительно малых энергетических потерях, однако эта операция, усложняя обработку, не всегда является необходимой. Так, если вещание ведётся в режиме «8К», то может осуществляться корреляционный обзор по задержке без режекции защитного интервала в диапазоне 0...895 мкс (до 134 км).

Структуры автокорреляционных функций сигнала ЦТВ в режимах «2К» и «8К» подобны и отличаются только масштабом по оси времени, который в режиме «8К» больше в 4 раза. Это связано с четырёхкратным отличием длительности символа. Таким образом, в рабочем диапазоне задержек коррелятора РЛС количество боковых пиков корреляционной функции в режиме «8К» оказывается меньше.

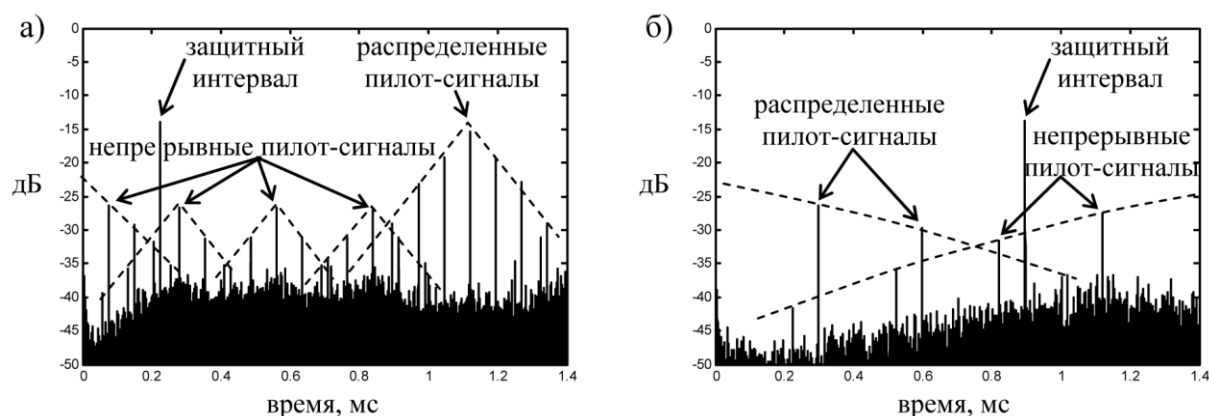


Рис. 6. Автокорреляционная функция сигнала цифрового ТВ:
а) в режиме «2К»; б) в режиме «8К»

Аналогично ФН сигнала цифрового ТВ в рабочем диапазоне доплеровских сдвигов и задержек коррелятора ПАРЛС имеет ряд побочных максимумов (рис. 7), координаты и амплитуды которых приведены в таблице 3.

Таблица 3

Амплитуды и координаты пиков ФН сигнала цифрового ТВ

Амплитуда, дБ	Время, мкс	Частота, Гц	Амплитуда, дБ	Время, мкс	Частота, Гц
в режиме «2К»			в режиме «8К»		
0	0	0	0	0	0
-13,9	224,0	0	-26,3	298,7	0
-26,4	74,7	0	-28,0	224,0	225
-26,7	280,0	0	-30,8	149,3	445
-27,7	56,0	895	-34,8	74,7	670
-29,1	149,3	0	-36,6	298,7	890

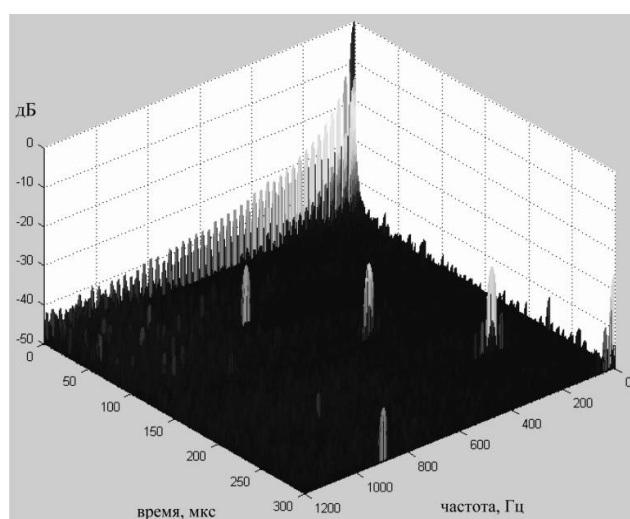


Рис. 7. Функция неопределённости сигнала цифрового ТВ в режиме «8К»

Стандартами цифрового радиовещания предусмотрена возможность использования нескольких видов модуляции ортогональных несущих: квадратурной фазовой манипуляции, равномерной и неравномерной 16- и 64-позиционной квадратурной амплитудной манипуляции. Однако установлено, что метод модуляции не оказывает заметного влияния на корреляционную функцию или ФН. Кроме того, данные характеристики обладают высокой стабильностью во времени.

3. Сравнительный анализ источников подсвета

Существующая сеть передатчиков аналогового ТВ и радиовещания и развёртываемая сеть передатчиков цифрового ТВ и радиовещания по пространственно-временному покрытию и диапазону частот удовлетворяют требованиям к источникам сигнала подсвета ПА РЛС. Сравнение существующих аналоговых и перспективных цифровых сигналов следует производить по таким важнейшим радиолокационным свойствам, как характеристики обнаружения и разрешения.

Поскольку все рассматриваемые сигналы являются непрерывными, их ФН однолепесткова по частоте и гарантирует однозначное определение доплеровской частоты сигнала с разрешающей способностью, обратной времени когерентного накопления. Характер ФН по задержке существенно различается. Из данных, приведённых в таблице 4, следует явное преимущество цифровых сигналов, которые обеспечивают лучшую на 1–2 порядка потенциальную разрешающую способность по дальности цели и лишены таких недостатков аналоговых сигналов, как нестабильность характеристик и неоднозначность определения дальности.

Сравнение характеристик обнаружения не может быть выполнено однозначно, так как зависит от мощностей передатчиков, условий распространения радиоволн, технических характеристик ПАРЛС. Поэтому приведём результаты анализа характеристик обнаружения, сделанного для конкретного телецентра, расположенного вблизи города Минска (табл. 5).

Таблица 4

Радиолокационные характеристики сигналов подсвета

Сигнал	Эффективная ширина спектра, кГц	Разрешение по дальности, м	Примечание
1	2	3	4
аналогового ТВ: – несущая сигнала изображения	–	–	обеспечивает определение только скорости цели
– сигнал яркости	100÷250	600÷1500	неоднозначное определение дальности с интервалом 9,6 км

Окончание табл. 4

1	2	3	4
– сигнал звукового сопровождения	10÷100	1500÷15000	высокая нестабильность характеристик
УКВ ЧМ радиовещания	10÷100	1500÷15000	высокая нестабильность характеристик
DAB	1536	98	–
DVB-T	7610	20	–

В данном телецентре для трансляции сигналов аналогового ТВ используются передатчики мощностью 5 кВт, из которой приблизительно 2/3 приходится на несущую и 1/3 на модуляционную составляющую сигнала яркости. Колебания полезной мощности, связанные с изменением в течение трансляции ширины спектра сигнала, количественно определяются коэффициентом q , указанным в таблице. Порог обнаружения определяет по отношению сигнал/помеха [4], при котором обеспечивается вероятность правильного обнаружения 0,8 и вероятность ложных тревог 10^{-2} . Запас по мощности представляет собой отношение полезной мощности передатчика к порогу обнаружения и определяет максимальную дальность обнаружения цели.

Таблица 5

**Сравнение качества обнаружения сигналов при подсвете
от телецентра в п. Колодищи**

Сигнал	Номинальная мощность передатчика, Вт	Фактор, определяющий полезную мощность	Полезная мощность передатчика, Вт (дБмВт)	Порог обнаружения, дБмВт	Запас по мощности, дБ
несущей изображения	5000	$p = 2/3$	3333 (65,2)	–126	191,2
яркости	5000	$p = 1/3;$ $q = 0.6...1$	1000...1667 (60...62,2)	–109	169...171, 2
ТВ звукового сопровождения	500	$q = 0.2...1$	100...500 (50...57)	–119	169...176
цифрового ТВ	1000	–	1000 (60)	–133	193

Как видно из таблицы 5, наибольшая дальность обнаружения обеспечивается при использовании сигнала цифрового ТВ, хотя мощность его передатчика в 5 раз ниже, чем у аналогового. Это объясняется низким уровнем боковых лепестков ФН сигнала цифрового ТВ.

Заключение

Передатчики аналогового ТВ и радиовещания ввиду их широкого распространения в настоящее время создают в пределах населённых территорий сплошное радиолокационное поле на малых высотах. В ряде стран зафиксирован положительный опыт создания, испытания и эксплуатации ПАРЛС, использующих данные сигналы. На основе анализа ФН показаны основные проблемы, связанные с использованием аналоговых сигналов, такие как нестабильность характеристик, неоднозначность определения времени запаздывания, ухудшение характеристик обнаружения в связи с высоким уровнем боковых лепестков ФН. Приемлемых характеристик ПАРЛС можно достичь путём многоканального построения и комплексирования обработки нескольких взаимодополняющих сигналов аналогового радиовещания.

В течение нескольких лет планируется полный переход на цифровые сигналы, такие как DVB-T, DAB, DRM. В связи с этим намечаются положительные перспективы для ПА радиолокации, наибольший интерес для которой представляют системы DVB-T и DAB, предназначенные для работы в диапазоне от 30 МГц до 3 ГГц. Анализ ФН сигналов цифрового радиовещания показывает, что они обладают высокой потенциальной разрешающей способностью по времени запаздывания, однозначностью оценок времени и частоты, стабильностью, а также низким уровнем боковых лепестков ФН, что обеспечивает более высокие характеристики обнаружения.

Список литературы

1. Использование ТВ-излучения для построения пассивного радиолокатора. *Le radar qui va "pirater" la television* // *Air et cosmos / Aviat. int.* – 1996. – Т. 33, № 1555. – С. 47. – Фр. (ВИНИТИ / РЖ: 24. Радиотехника. 24В. Радиолокация. Радионавигация. Радиоуправление. Телевизионная техника (Выпуск сводного тома). – 1996. – № 8. – 8В20).
2. Howland P.E. *Television based bistatic radar: PhD Thesis, School of Electronics and Electrical Engineering, University of Birmingham, UK, 1997.*
3. Пат. 6319, МПК7 G 01S 13/00. Полуактивная радиолокационная станция / Охрименко А.Е., Курлович В.И., Папушой В.И., Шаляпин С.В. [и др.]. – № a20000555; Заявл. 13.06.2000; Оpubл. 30.12.2001 // Официальный бюллетень / Национальный центр интеллектуальной собственности Республики Беларусь. – 2001. – № 4. – С. 67.

4. Охрименко А.Е. Основы обработки и передачи информации. – Минск: МВИЗРУ ПВО, 1990. – 180 с.
5. Охрименко А.Е. Первенец белорусской радиолокации – скрытный радар «Поле»: научно-технический очерк. – Минск: БГУИР, 2005. – 167 с.
6. Griffiths H.D. From a different perspective: principles, practice and potential of bistatic radar // Proc. of the International Radar Conference, Adelaide, Australia, 3–5 September 2003. – P. 1–7.
7. Envisat overview [Electronic resource] / European Space Agency. – 2005. – Mode of access: http://www.esa.int/esaEO/SEMWYN2VQUD_index_0_m.html.
8. Методы радиолокационного распознавания и их моделирование / Я.Д. Ширман, С.А. Горшков, С.П. Лещенко [и др.] // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники. – 1996. – № 11.
9. Семашко П.Г. Корреляционное различение гипотез однозначной дальности в РЛС с телевизионным подсветом // Радиолокация, навигация, связь: материалы IX Междун. науч.-технич. конф., 22–24 апреля 2003 г., Воронеж. – Воронеж, 2003. – Т. 3. – С. 1393–1403.
10. Семашко П.Г., Охрименко А.Е. Характеристики обнаружения телевизионного сигнала яркости в радиолокационной системе с посторонним подсветом // Доклады БГУИР. – 2005. – № 2(10). – С. 60–64.
11. Попов И.В. Освещение воздушной обстановки с использованием излучения вещательных станций КВ диапазона / И.В. Попов, И.В. Луценко, В.И. Луценко / под ред. А.И. Громыко, А.В. Сарафанова // Современные проблемы радиоэлектроники: сб. науч. тр. – М.: Радио и связь, 2006. – С. 25–28.
12. Луценко И.В., Попов И.В., Луценко В.И. Бистатические РЛС с подсветкой ионосферными сигналами связных станций коротковолнового диапазона // Радиофизика и электроника. – 2007. – Т. 12, № 1. – С. 199–203.
13. Silent Sentry. Innovative Technology for Passive, Persistent Surveillance [Electronic resource] / Lockheed Martin Corp. – 2005. – URL: <http://www.lockheedmartin.com/data/assets/10644.pdf>.
14. Ксендзук А.В., Фатеев В.Ф., Попов С.А. Неизлучающая радиолокационная система, основанная на приёме отражённых сигналов навигационных систем ГЛОНАСС и GPS // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники. – 2009. – № 4. – С. 60–65.

15. European Standard ETSI EN 300 401 V1.3.3 (2001-05) «Radio Broadcasting Systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers» [Electronic resource] / European Telecommunications Standards Institute. – 2001. – URL: <http://www.etsi.org>.
16. European Standard ETSI EN 300 744 V1.4.1 (2001-01) «Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television» [Electronic resource] / European Telecommunications Standards Institute. – 2001. – URL: <http://www.etsi.org>.
17. Семашко П.Г., Охрименко А.Е. Скрытная радиолокация с использованием сигналов цифрового телевизионного вещания // Доклады НАНБ. – 2005. – № 6. – С. 37–41.
18. Saini R., Cherniakov M. Investigation of digital TV-terrestrial signal for radar application // Proc. of the XI European signal processing conference (EUSIPCO), Toulouse, France, 3–6 Sept. 2002. – Vol. 3. – P. 15–18.

РАЗДЕЛ 6

WEB-ТЕХНОЛОГИИ

УДК 004.773

Е.А. Бахман^{*}, А.В. Кизим^{**}

АВТОМАТИЗАЦИЯ ПРИКЛАДНЫХ ЗАДАЧ С ПОМОЩЬЮ WEB-ОРИЕНТИРОВАННЫХ СИСТЕМ

В данной работе рассматривается задача повышения эффективности работы пользователей (в частности, при поиске информации) в сети Интернет за счёт автоматизации прикладных задач с помощью web-ориентированных систем.

В рамках работы спроектирован и разработан web-портал «Здоровый город», посвящённый вопросам здоровья и психологии, а также реализован программный модуль для осуществления миграции web-сайтов из CMS Joomla в CMS Drupal.

На современном этапе развития информационных технологий сеть Интернет является одним из главных источников информации. Однако достаточно часто из-за многообразия и недостаточной систематизации интернет-ресурсов доступ пользователей к необходимой информации затруднён. В результате значительно снижается эффективность работы конечных пользователей, возникает необходимость дополнительно затрачивать время и ресурсы на систематизацию полученной информации.

В качестве предметной области были рассмотрены медицина, здравоохранение, психология, личностный рост и саморазвитие. Выбор этой тематики неслучаен: потребность пользователей в получении информации о здоровье, психологии, системе здравоохранения достаточно велика. Региональная популярность данной тематики (среди жителей Волгограда и Волгоградской области) выше средней на 6,5 %. В данной работе был проведён анализ потребностей пользователей по тематике «Здоровье» и «Психология». При анализе рассматривались запросы пользователей сети Интернет в поисковых системах Rambler, Yandex. Согласно данным сервиса «Статистика ключевых слов» (<http://wordstat.yandex.ru>), слово «лечение» пользователи запрашивали около 2 787 640 раз в месяц, а «здоровье» и «психология» – 924 278 и 1 436 350 раз соответственно.

^{*} 400131, Волгоград, пр. Ленина, 28, ВолгГТУ, vega0707@rambler.ru

^{**} 400131, Волгоград, пр. Ленина, 28, ВолгГТУ, kizim@mail.ru

Цель данной работы заключается в повышении эффективности работы пользователей сети Интернет (в частности, сокращении временных затрат (например, при поиске информации) и улучшении эргономичности). Под «эргономичностью» понимается удобство работы с web-ресурсом, которое подразумевает прежде всего систематизацию информации, предоставляемой пользователю; организацию быстрого и удобного доступа к информации.

Для достижения поставленной цели решались следующие задачи:

- анализ потребностей пользователей сети Интернет по тематике «Здоровье» и «Психология»;
- анализ программных средств, которые позволяют автоматизировать решение прикладных задач – систем управления контентом (CMS);
- проектирование и реализация web-портала «Здоровый город» по вопросам медицины, здравоохранения, психологии и саморазвития.

При разработке web-ресурса возникла необходимость переноса сайта из одной CMS в другую, поэтому в рамках данной работы также рассмотрена задача автоматизации переноса (миграции) сайта из одной CMS в другую, решение которой также позволяет повысить эффективность работы пользователей сети Интернет и web-разработчиков, т.к. технические характеристики использования системы в значительной степени влияют на скорость загрузки web-ресурса.

Средством повышения эффективности работы конечных пользователей является автоматизация прикладных задач с помощью web-ориентированных систем – систем управления контентом (CMS). Это программный комплекс, позволяющий автоматизировать процесс управления сайтом (информационным наполнением, структурой, правами доступа и др.), а также предоставляющий дополнительные функциональные возможности, такие как рассылка, статистика, поиск, средства коммуникации с пользователями и др.

В данной работе были рассмотрены исключительно открытые CMS (Drupal, Joomla, MODx, TYPO3), проведён их сравнительный анализ. Основные параметры сравнения – системные требования, безопасность, поддержка, удобство использования, гибкость системы и производительность, функциональность [1].

Рассмотрим основные этапы проектирования и разработки web-портала «Здоровый город»:

Определение цели и задач сайта, выявление целевой аудитории.

Цель портала «Здоровый город» была сформулирована [2] как повышение информационной осведомлённости жителей города в вопросах здоровья и психологии, улучшение качества медицинских и прочих сопутствующих услуг в сфере здравоохранения.

Задачи портала «Здоровый город»:

- предоставление пользователям информации по наиболее актуальным вопросам здоровья, психологии, саморазвития и о здравоохранении в Волгограде;
- систематизация информации и каталогизация ресурсов по указанной тематике;
- пропаганда профилактики заболеваний и здорового образа жизни.

Основное отличие разрабатываемого сайта – взаимосвязь информации о здоровье и психологии, рассмотрение актуальных вопросов медицины и здравоохранения в регионе (например, поиск информации о медицинских учреждениях Волгограда и Волгоградской области).

Логическое проектирование. На данном этапе были спроектированы структуры сайта и система навигации.

Информация на сайте сгруппирована по тематическим разделам:

- «Здоровье и красота» (здоровый образ жизни, правильное питание, народные советы, новости медицины, диагностика и лечение, секреты красоты);
- «Психология» (общие понятия психологии, человек и общество, отношения, советы психолога, взрослым о детях, саморазвитие и успех);
- «Жизнелогия» (философия мысли, мировосприятие);
- «Здравоохранение» (здравоохранение как система, медицинские учреждения).

Система навигации сайта представлена в виде верхнего и главного меню. Верхнее меню состоит из пунктов «О проекте», «Карта сайта», «Контактная информация», «Обратная связь», главное меню – из пунктов «Здоровье и красота», «Психология», «Здравоохранение», «Форум», «Полезные ссылки».

Физическое проектирование. На данном этапе были выбраны технологии и ПО, которые должны быть использованы при разработке web-портала «Здоровый город»:

- в качестве ядра была выбрана CMSJoomla, которая отличается удобством использования, гибкостью, легко расширяется с помощью дополнительных модулей и плагинов [3]. Поскольку возникла потребность переноса сайта из-за превышения лимитов ресурсов хостинга, для дальнейшей разработки web-портала была выбрана CMSDrupal. Отличительные особенности Drupal – высокая производительность, мощная функциональность, оптимизация системы (сжатие и кэширование страниц, наличие механизма отключения функций при перенагрузке) [1];
- разработка дизайна сайта;
- настройка функционала;

- на сайте функционируют следующие сервисы: поиск информации – обычный и расширенный, облако тегов, обратная связь, карта сайта, авторизация пользователей (в том числе регистрация новых пользователей и восстановление пароля для зарегистрированных пользователей), опросы (голосования), подписка на рассылку материалов;
- наполнение контентом;
- размещены 120 статей по тематическим рубрикам: «Здоровье и красота», «Психология» и «Здравоохранение», а также информация о проекте и контактная информация.
- комплексная отладка и тестирование сайта;
- выполнено тестирование сайта на наличие ошибок при выполнении различных запросов пользователей, на правильность отображения материала в различных веб-браузерах: InternetExplorer, MozillaFirefox, Opera.
- публикация в сети Интернет.
- сайт опубликован в сети Интернет. Доступ к нему осуществляется по адресу: <http://www.ty-sam.ru>.

Поскольку в данной работе была рассмотрена задача автоматизации миграции сайта, также был спроектирован и реализован программный модуль миграции из Joomla в Drupal, который в отличие от существующих решений позволяет переносить в Drupal не только контент (статьи, разделы, категории) и пользователей, но и голосования, комментарии к статьям [4]. В этом заключаются техническая новизна и практическая значимость данной разработки.

Таким образом, для повышения эффективности работы пользователей сети Интернет, в частности для сокращения временных затрат:

- проведён анализ CMS как средства для разработки портала о здоровье и психологии «Здоровый город»;
- спроектирован и реализован web-портал «Здоровый город»;
- спроектирован и реализован программный модуль для автоматизации миграции из CMS Joomla в CMS Drupal;
- достигнуто повышение эффективности работы пользователей сети Интернет за счёт сокращения временных затрат при поиске информации о здоровье и психологии;
- достигнуто повышение производительности сервера хостинга за счёт миграции.

В настоящее время на сайте опубликовано более 120 статей, зарегистрировано около 60 пользователей, 145 пользователей оформили рассылку на новости сайта, что подтверждает практическую и социальную значимость проекта.

Перспектива дальнейшей работы связана с использованием web-ориентированных систем в автоматизации прикладных задач в области электронной коммерции.

Основное направление для дальнейшего развития портала «Здоровый город» – развитие форума и тематического раздела «Здравоохранение». Поэтому очень важным является привлечение к сотрудничеству медицинских учреждений, врачей и специалистов, а также всех заинтересованных лиц.

Список литературы

1. CMS Matrix. – URL: <http://www.cmsmatrix.org>.
2. Вандюк Д. CMSDrupal. Руководство по разработке системы управления сайтом: пер. с англ. – М. : Вильмс, 2008.
3. Рамел Д. Самоучитель Joomla: пер. с англ.. – СПб.: БХВ-Петербург, 2008.
4. Обоснование потребности и постановка задач Веб-портала «Здоровый Волгоград» / Е.А. Бахман, А.В. Кизим // Информационные технологии в образовании, технике и медицине : матер. междунар. конф., 21–24 сент. 2009 / ВолгГТУ [и др.]. – Волгоград, 2009. – С. 107.
5. Савитский В. Миграция из Joomla в Drupal – URL: <http://drupalcookbook.ru/recept/migracija-joomla-drupal>.

РАЗДЕЛ 7

МЕНЕДЖМЕНТ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

УДК 658:004

В.И. Божич, М.Б. Савченко, З.И. Попова

КОНЦЕПТУАЛЬНАЯ МОДЕЛЬ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ СОВМЕСТНОЙ ДЕЯТЕЛЬНОСТИ

Рассматриваются особенности совместной деятельности естественного и искусственного интеллектов с учётом ментальных и интенциональных характеристик на основе мультиагентной парадигмы.

Введение. Влияние и значимость человеческого фактора на процесс принятия решений накладывает отпечаток на развитие современной теории принятия решений, принципиальным моментом которого является снятие противопоставления между «исследователем» и «объектом исследования» [1]. Проблема учёта активности лица, принимающего решение, становится первичной и требует отступления от концепции «нормативных моделей управления» к моделям «совместной деятельности».

Создание информационно-психологической системы совместной деятельности связано с разработкой двух взаимосвязанных моделей – естественного и искусственного интеллектов. Достоверность таких моделей позволит целенаправленно вести процесс управления на различных уровнях его представления и обеспечивать адаптивную настройку и подстройку тех или иных информационных процессов в соответствии с индивидуальными характеристиками конкретной личности.

Значимость информационно-психологической деятельности. Перенос принципов человеческой психики на компьютерные модели не может быть совершён простыми и формальными средствами. Для интерпретации деятельности человека должны создаваться интеллектуальные модели, которые могли бы учитывать их закон функционирования в зависимости от многих факторов, главным из которых является фактор учёта индивидуальных психофизиологических особенностей человека.

Современные научно-технические достижения должны обеспечить в ближайшем будущем естественный человеко-машинный интерфейс [2, 3]. В настоящее время имеются предпосылки для разработки такого интерфейса:

- новые особенности компьютеров (мультимедиа) позволили реструктурировать представление знаний в компьютерах и тем самым расширили формы представления и обработки информации;
- новые особенности нейрокомпьютинга позволили создавать системы, которые могут настраиваться и перестраиваться на различные законы функционирования в результате применения процедур обучения, что характерно для систем живой природы;
- новые алгоритмы эволюционного поиска и алгоритмы генетической оптимизации, по аналогии с живой природой, позволяют вести целенаправленный поиск разумных стратегий поведения;
- новые модели мультиагентных и бионических систем «искусственная жизнь» позволяют работать в пространстве поиска «большой размерности».

Современные результаты когнитивной психологии и искусственного интеллекта позволяют ставить и решать новые задачи построения интегрированных компьютерных систем «совместной деятельности» [4, 5]. Такие системы должны реализовывать процедуры, наделённые психикой деятельности. При этом важным вопросом является выбор технологических средств искусственного интеллекта, которые будут удовлетворять целостности объектов, порождаемых естественным и искусственным интеллектами.

Возможные пути решения задач по выделению указанных отношений должны быть основаны на симбиозе образного и символично-логического представления знаний и их обработки в системах искусственного интеллекта (ИИ), позволяющих работать с образами так же оперативно, как с символами.

Примером тому являются когнитивные карты и сети, представляющие собой интеллектуальные средства генерации решений путём «фокусировки внимания» образного мышления лица, принимающего решение (ЛПР) [1]. По утверждениям психологов, визуализация представления информации является инструментом, позволяющим ЛПР ускорить и повысить эффективность решений [6, 7].

Акт формирования внимания ЛПР является актом интерпретации образа и, следовательно, рассуждением, обобщением, формированием гипотез. Использование образного мышления ЛПР в рамках совместной деятельности с системой ИИ является механизмом ускоренного мышления на основе использования когнитивных свойств информации.

В этой связи разработка новых технологий принятия решений должна учитывать результаты агентно-ориентированного подхода [8], который позволяет учесть представление ментальных и интенциональных моделей деятельности ЛПР и системы ИИ в форме нечётких понятий и правил лингвистического и когнитивного характера.

Использование результатов такого подхода позволит учесть значимость знаковых представлений семиотики с точки зрения формирования взаимодействия образного мышления ЛПР и символьного мышления системы ИИ. Аспект семиотического рассмотрения образа позволит трансформировать проблему значимости образа в проблему координации деятельности.

Такая оценка значимости образа позволит использовать концептуальные модели в виде наглядного представления семантических отношений «совместной деятельности» на уровнях образного и символьного мышлений «агент – действие – объект».

В настоящее время большинство пользователей систем когнитивного моделирования – это финансовые и политические аналитики, ситуационные центры военного и политического руководства развитых стран. Развитию когнитивных нечётких схем активно уделяют внимание промышленные гиганты (Motorola, General Electric, Otis Elevator, Pacific Gas & Electric, Ford и другие). Такие программы – это сложные комплексные системы, требующие определённых усилий по разработке, освоению и настройке. Они являются закрытой информацией.

Создание *мультиагентных когнитивных карт* как активных интеллектуальных средств оценки ситуации и решения проблем информационно-психологической безопасности позволяет связать воедино виртуальное и реальное пространства, психологию естественного и искусственного интеллектов, символьное компьютерное мышление и образное мышление человека.

Модель когнитивного агента. Построение концептуальной модели совместной деятельности предопределяет использование когнитивных агентов, которые имеют систему ментальных ценностей и определяют варианты поведения в рамках *семантической конструкции* «агент – действие – объект».

Модель когнитивного агента «совместной деятельности» представляет собой тезаурус с размеченными семантическими связями. При этом «знания и убеждения», а также «цели и намерения» можно отнести к категории нечёткого отношения между состоянием агента и предполагаемым результатом его деятельности.

С помощью указанных функций поведения когнитивного агента можно описать процесс подготовки решения с учётом ментальных характеристик. В целом концептуальная модель когнитивного агента представляет собой тезаурус с размеченными семантическими связями – множество понятий и отношения между ними:

$$T(NET) = (G, P_M, R, \lambda, \eta),$$

где $G = G_1 \cdot \dots \cdot G_n$ – семейство понятий для различных подграфов сети; $P_M \subset G$ – множество ментальных понятий и суждений; $\lambda: G \rightarrow [0, 1]$ – функция принадлежности, которая упорядочивает понятия; $\eta: R \rightarrow [0, 1]$ – функция принадлежности, которая описывает веса связей.

При таком подходе к представлению семантической информации когнитивный агент косвенно учитывает информацию путём изменения его тезауруса, что даёт возможность говорить об обобщающем тезаурусе:

$$T(NET) = T(NE_{T_1}) \cap \dots \cap T(NE_{T_k}),$$

где $T(NE_{T_i})$ – подграф исходного образа сети.

Для реализации возможности семиозиса знаков с учётом интенциональных отношений следует использовать М-сети [9], так как для формирования множества мотивов и стремлений ментальная подсистема когнитивных агентов должна содержать механизмы обработки знаний и планирования действий. При этом «знания и убеждения» и «цели и намерения» можно отнести к категории нечёткого отношения моделирования – нечёткого соответствия между состоянием агента и предполагаемым результатом его деятельности.

Содержание убеждений и знаний рассматривается в форме лингвистических предложений, а содержание намерения рассматривается как действие в рамках норм и обязательств, которые их ограничивают. Для формирования множества мотивов и стремлений ментальная подсистема когнитивных агентов должна содержать механизмы обработки знаний и планирования действий. Соотношение между побуждением и действием определяет нечёткое описание действий агента. Деятельность агента связана:

- с формированием желаний $Ж = T(G_i) \cdot P_M$;
- выполнением намерений (действий) $H = T(G_i) \rightarrow T(G_S)$;
- пересмотром убеждений $Y = T(G_i)$.

В связи с наличием у агента шкалы собственных предпочтений, которая не может охватить моделируемый мир, целесообразно интенциональные характеристики агента (показатели уверенности) также описывать на основе нечётких множеств. Тогда состояние агента в любой момент времени характеризуется тройкой $(Ж, H, Y)$.

Пусть $F(\bullet)$ – класс нечётких множеств на множестве $(\bullet)$. Тогда функция пересмотра убеждений агента может быть задана на основе текущего восприятия и текущих убеждений формированием нового нечёткого множества убеждений:

$$g: F(B) \cdot F(Y) \rightarrow F(Y).$$

Функция формирования выбора q может быть записана как отображение декартового произведения множества убеждений и намерений на множество желаний:

$$q: F(Y) \cdot F(H) \rightarrow F(Ж).$$

В рамках «доски объявления совместной деятельности» [8] оператор текущего восприятия определяет нечёткое отношение процесса моделирования семиозиса знаков как со стороны агента, так и со стороны ЛПР:

$$B: T(G_i) \cdot P_{MA} \cdot P_{ЛПР} \rightarrow [0, 1],$$

где P_{MA} – ментальные представления агента; $P_{ЛПР}$ – ментальные представления ЛПР на уровне образного мышления.

С помощью указанных функций можно описать процесс рекурсивного уточнения некоторой иерархической структуры плана действий с учётом ментальных характеристик, благодаря фокусировке внимания ЛПР с помощью интеллектуальных инструментальных средств когнитивных агентов.

Отличительной особенностью предлагаемого подхода является *интеграция нейросетевой модели*, которая имитирует характеристики объекта в виде отношения «вход – выход», и *семантической сети*, отражающей внутреннюю детализацию объектов и их отношений. Предлагается развитие концепции М-автоматов и сетей для решения задачи, направленной на «включение» образного мышления ЛПР в контур «объект управления – система управления».

Семантическая М-сеть является статической моделью, отражающей совокупность объектов (*i-моделей*) и их взаимоотношения. Способ представления информации через М-сеть носит отпечатки классической парадигмы моделирования путём декомпозиции сложного объекта на части и установления взаимосвязи каждой части объекта с моделью нейрона или нейронного ансамбля.

Такое представление нейронной сети (НС) в виде семантического графа фактически является *процессом когнитивного обучения нейронной сети*, который по своей сути отличается от процесса классического обучения НС с помощью обучающих выборок. Когнитивное обучение определяется двумя последовательными процедурами:

- обучение через формирование множества объектов и установление их значимости;
- обучение через формирование множества отношений между объектами путём задания весов связей между объектами. Вторая процедура когнитивного обучения по своей сути, но не по процедуре, напоминает обучение классических НС.

Каждому объекту М-сети соответствует нечёткое понятие, определённое через экспертные оценки и заданные значениями весов синапсов и порогов *i-моделей*. Данное обстоятельство позволяет рассматривать модель М-сети как НС по сути построения, но нечёткую систему по процессу функционирования. При этом выходной слой НС (агрегатор и дефузификатор) обобщает нечёткую информацию, поступающую от нейронов.

Параметры всех нейронов-объектов среднего слоя являются нечёткими понятиями, функция принадлежности которых определяется системой «усиления-торможения», управляющей процессом возбуждения *i*-моделей: если возбуждённость элемента выше некоторого заданного параметра или выше возбуждённости других элементов, то порог такого элемента уменьшается.

В результате исходный семантический граф М-сети можно представить подграфом, состоящим из вершин, которые имеют повышенную возбуждённость, и связей между ними. Если отобразить возле каждой *i*-модели значимость её возбуждения с помощью цифры, цвета или другого визуального способа, то получим когнитивный граф-образ. Представление М-сети как когнитивной карты позволяет осуществлять манипуляции со смыслом символов: вершины соответствуют некоторым объектам, связи определяют отношения между объектами.

Система «усиления-торможения», воздействуя на нейронную М-сеть, каждый раз представляет ЛППР наиболее активную когнитивную информацию и динамику её изменения. Визуальный анализ такого графа позволяет ЛППР учитывать на сознательном или подсознательном уровнях сложившуюся ситуацию, оценить её и принять соответствующие решения и действия.

В результате ЛППР имеет возможность управлять детализацией просмотра когнитивного графа в зависимости от его целей и особенностей восприятия информации. Таким образом, предлагаемая модель позволяет соотнести (связать, объединить) действия ЛППР и системы искусственного интеллекта и тем самым соединить субъективные и объективные не только декларативные, но и процедурные знания.

Для настройки параметров когнитивной нейронной сети предлагается использовать механизм настройки классических НС. Такой механизм рассматривает НС как модель, которая отражает только поведение моделируемого объекта и не учитывает внутреннюю детализацию (семантику) нейронной сети.

Процедура формирования такой модели фактически является процедурой когнитивного сжатия информации об объекте через обучение нейронной сети в соответствии с заданной обучающей выборкой. При этом задача обучения разбивается на две части:

- поиск набора обучающей выборки, что можно выполнить путём анализа отношений «объект управления – система управления»;
- применением процедуры настройки когнитивной нейронной сети путём предъявления набора обучающих выборок по аналогии с обучением нейронных рекуррентных или рециркуляционных сетей [10].

Процесс настройки параметров когнитивной нейронной сети должен содержать этап предварительного экспериментального тестирования, а также должен учитывать фактор учёта субъективности ЛПР. Семантический граф М-сети, определяющий ситуацию (сценарий), может иметь вершины-объекты (*i*-модели), отражающие ментальные и интенциональные характеристики (грусть, апатия, радость и т.д.), что позволяет контролировать психофизиологическое состояние ЛПР в процессе работы с когнитивной нейро-нечёткой моделью.

Фокусировка внимания и семантические сети. Для фокусировки внимания естественного и искусственного интеллектов с учётом ментальных и интенциональных характеристик совместной деятельности предлагается использовать гибридную нейронную модель на основе совмещённого синтеза М-сети и АРТ-сети, которая является искусственным когнитивным анализатором образов и реализует алгоритм кластеризации.

Образцом первого кластера считается первый входной вектор; все последующие входные векторы сравниваются с образцом первого кластера. Для всех последующих входных векторов происходит измерение значимости текущего входного вектора с имеющимися образцами кластеров. Результатом измерения является либо новый кластер и соответствующий ему образец, либо отнесение входного вектора к тому или иному кластеру. В качестве метрики может служить некоторая заранее задаваемая величина расстояния между векторами в виде порога.

В целом когнитивный анализатор в виде АРТ-сети [11] определяет множество абстрактных образов – образцов кластеров: входной вектор – набор признаков абстрактного образа, значения порогов – мера определения множества таких абстрактных образов. Таким образом, нейронная АРТ-сеть может выполнять дополнительные функции системы «торможения-усиления» М-сети, связанные с искусственным когнитивным анализом как отдельных вершин семантического графа М-сети, так и подграфов, выделенных в процессе функционирования М-сети.

Добавление или исключение образов или понятий в семантическом графе возможно реализовать путём корректировки (усреднения) весов нейронов в зависимости от желаемых выходов. Каждый вновь рассматриваемый образ может создавать дополнительные классификационные категории, оставляя неизменными результаты предшествующего обучения (стабильность и пластичность гибридной нейронной сети).

Известно, что геометрическая интерпретация распознавания может быть трансформирована в лингвистический подход с помощью выделения набора типичных фрагментов образов и характеристик взаимного расположения фрагментов (словаря и правил).

Такой подход имеет «когнитивную аналогию» и требует не столько знаний основ математической лингвистики, сколько специальных приёмов для лингвистической обработки образов. Заданием пространства признаков формулируется свойство компактности для разделения абстрактных образов на основе абстрактного пространства. Если входной вектор-образ не соответствует ни одному из ранее полученных образов, то создаётся новая категория для запоминания образа.

Заметим значимую особенность: АРТ-сеть может распознать образ, никогда не виденный ранее. Наличие эволюционных процедур формирования нейронной сети и системы «усиления-торможения» М-сети приводит к важному достоинству – *генерации факта существования заранее неизвестной информации* в заданном пространстве образов, о которых ЛПП ранее не имел никакого представления.

Данное обстоятельство может быть учтено при описании принципов эволюции и роста семантического графа с целью визуализации представления информации и «вмешательства» естественного когнитивного мышления ЛПП. В результате обучение и самообучение естественного интеллекта и(или) системы ИИ могут служить мерой пригодности (мерой оценки) выбранного пространства образов.

Степень совпадения абстрактных образов, выделяемых в процессе такого самообучения относительно реальных образов, может служить оценкой правильности выбора абстрактного пространства. Оценка деятельности через визуализацию некоторых образов позволяет создавать систему управления на позициях когнитивной психологии при решении задач формирования внимания ЛПП.

Заключение

Интеграция информационной и психологической сфер влияния приводит к появлению новых форм взаимодействия естественного и искусственного интеллектов. Результатом этого должны стать интеллектуальные инструментальные средства и ресурсы, обеспечивающие выполнение функций управления запросами, анализ эвристик и стратегий, производства гипотез, интегрированных концепций и других функций, которые будут появляться по мере развития указанной интеграции.

Такие инструментальные средства должны нести в себе функции человеко-машинных интерфейсов, систем поддержки принятия решений, систем искусственного интеллекта и в то же время быть ориентированы в первую очередь на естественный интеллект, его запросы и действия, защищать физическое, умственное и психологическое здоровье человека.

Список литературы

1. Трахтенгерц Э.А. Компьютерная поддержка принятия решений: научн.-практич. издание // Серия «Информатизация России на пороге XXI века». – М.: СИНТЕГ, 1998. – 376 с.
2. Люггер Джордж Ф. Искусственный интеллект: стратегии и методы решения сложных проблем: пер. с англ. – 4-е изд. – М.: Вильямс, 2003. – 864 с.: ил.
3. Рассел С., Норинг П. Искусственный интеллект: современный подход: пер. с англ. – 2-е изд. – М.: Вильямс, 2006. – 1408 с.: ил.
4. Солсо Р. Когнитивная психология. – СПб.: Питер, 2002. – 592 с.: ил. – (Серия «Мастера психологии»).
5. Андерсон Дж. Когнитивная психология. – 5-е изд. – СПб.: Питер, 2002. – 496 с.: ил. – (Серия «Мастера психологии»).
6. Дормашев Ю.Б., Романов В.Я. Психология внимания: учебник. – 3-е изд., испр. – М.: Московский психолого-социальный институт: Флинта, 2002. – 376 с.
7. Зинченко Т.П. Память в экспериментальной и когнитивной психологии. – СПб.: Питер, 2002. – 320 с.: ил. – (Серия «Мастера психологии»).
8. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 352 с. (Науки об искусственном).
9. Автоматы и разумное поведение. Опыт моделирования / Н.М. Амосов, А.М. Касаткин, Л.М. Касаткина, С.А. Талаев. – Киев: Изд-во «Наукова думка», 1973. – 374 с.
10. Хайкин С. Нейронные сети: полный курс: пер. с англ. – 2-е изд. – М.: Вильямс, 2006. – 1104 с.: ил.
11. Гроссберг С. Внимательный мозг // Открытые системы. – 1997. – № 4 – С. 29–33.

УДК 658.3:331

М.Б. Савченко

ИНФОРМАЦИОННО-ПСИХОЛОГИЧЕСКАЯ БЕЗОПАСНОСТЬ «БЕЛЫХ ВОРОТНИЧКОВ»

В работе предлагается анализ соответствия современной оценки умственного труда условиям информационного общества.

Введение. С целью разделения профессий *сферы материального производства и сферы интеллектуальных услуг* индустриальное общество выделило две категории работников, называемых «синими воротничками» и

«белыми воротничками». Границы такого разделения профессий в постиндустриальном обществе стали размытыми в связи с автоматизацией рабочей деятельности.

Дальнейшее повышение эффективности труда связано с информатизацией большинства сфер деятельности и проявлением для различных профессий характерных особенностей *умственного труда* «белых воротничков». Внедрение высоких технологий в административные, промышленные и сервисные сферы изменили потребность в навыках работников, которые ориентированы на выполнение одиночных задач. Сложные технологии привели к распределённым системам управления, диагностики и контроля информационных процессов.

В информационном обществе формы интеллектуального труда характерны как для профессий *сферы материального производства*, так и для профессий *сферы интеллектуальных услуг*. Число людей, занятых в сфере интеллектуальных услуг, неукоснительно увеличивается. Этот факт неизбежен, так как переход общества к постиндустриальной и информационной стадиям развития определён социально-экономическим критерием:

- если в обществе более 50 % населения занято в сфере услуг, наступила постиндустриальная фаза его развития;
- если в обществе более 50 % населения занято в сфере информационно-интеллектуальных услуг, общество становится информационным.

Информатизация трудовой деятельности и повседневной жизни людей (вне рабочей среды) усиливает нагрузки, свойственные умственному труду. Компьютерные технологии привели к проблеме информационно-психологической безопасности личности. Следствием этому являются процессы формирования сердечно-сосудистой патологии, угнетения психики, ухудшения функций внимания, памяти, восприятия у большинства людей. В этих условиях актуальным является анализ состояния дел по оценке современным обществом понятия «*профессиональный риск работников умственного труда*».

1. Различение форм умственного труда

В настоящее время общепризнанной является *физиологическая классификация трудовой умственной деятельности*, которая основана на различении форм труда:

- формы труда, связанные с *управлением производственными процессами и механизмами*. Процесс управления полуавтоматическим и автоматическим производством включает человека в систему как необходимое оперативное звено. Как правило, оператор находится в состоянии непрерывной напряжённой готовности к действию;

- труд руководителей учреждений, предприятий (*управленческий труд*) характеризуется чрезмерным ростом объёма информации, возрастанием дефицита времени для её переработки, повышенной личной ответственностью за принятое решение, периодическим возникновением конфликтных ситуаций;
- труд *преподавателей и медицинских работников* отличается постоянными контактами с людьми, повышенной ответственностью, часто дефицитом времени и информации для принятия правильного решения, что обуславливает высокую степень нервно-эмоционального напряжения;
- труд *учащихся и студентов* характеризуется напряжением основных психических функций (память, внимание, восприятие) и наличием стрессовых ситуаций (экзамены, зачёты);
- наиболее сложная форма трудовой деятельности, требующая значительного объёма памяти, напряжения, внимания, – это *творческий труд*. Труд научных работников, конструкторов, писателей, архитекторов приводит к значительному повышению степени нервно-эмоционального напряжения. При этом можно наблюдать тахикардию, повышение кровяного давления, изменения ЭКГ, увеличение лёгочной вентиляции и потребления кислорода, повышение температуры тела и другие сдвиги со стороны вегетативных функций.

2. Показатели напряжённости интеллектуального процесса

Умственный труд объединяет работы, связанные с приёмом и передачей информации, требующие преимущественного напряжения сенсорного аппарата, внимания, памяти, активизации процесса мышления. Основным показателем умственного труда является напряжённость труда – характеристика трудового процесса, отражающая преимущественную нагрузку на центральную нервную систему. Для профессиональных групп исполнительского, управленческого, операторского и творческого видов труда используются следующие показатели:

- *нагрузки интеллектуального характера:*
 - а) содержание работы;
 - б) восприятие сигналов (информации) и их оценка;
 - в) распределение функций по степени сложности задания;
 - г) характер выполняемой работы;
- *сенсорные нагрузки:*
 - а) длительность сосредоточенного наблюдения;
 - б) плотность сигналов (световых, звуковых) и сообщений в среднем за 1 час работы;
 - в) число производственных объектов одновременного наблюдения;

- г) размер объекта различения при длительности сосредоточенного внимания;
- д) работа с оптическими приборами при длительности сосредоточенного наблюдения;
- е) наблюдение за экраном видеотерминала;
- ж) нагрузка на слуховой анализатор;
- з) нагрузка на голосовой аппарат;
- *эмоциональные нагрузки*:
 - а) степень ответственности за результат собственной деятельности, значимость ошибки;
 - б) степень риска для собственной жизни;
 - в) степень ответственности за безопасность других лиц;
- *монотонность нагрузок*:
 - а) число элементов (приёмов), необходимых для реализации простого задания или многократно повторяющихся операций;
 - б) продолжительность выполнения простых производственных заданий или повторяющихся операций;
 - в) время активных действий;
 - г) монотонность производственной обстановки;
- *режим работы*:
 - а) фактическая продолжительность рабочего дня;
 - б) сменность работы;
 - в) наличие регламентированных перерывов и их продолжительность (без обеденного перерыва).

3. Руководство гигиенической классификацией Р 2.2.755–99

Установка на защитное поведение в интеллектуальном труде зависит от трёх факторов:

- 1) степени предполагаемого риска;
- 2) преобладающей мотивации;
- 3) опыта работы.

Тем самым обладание информацией о существующем риске даёт важный рычаг управления безопасностью трудовой деятельности.

В качестве основы метода оценки индивидуального профессионального риска для количественной характеристики условий труда наиболее полные и точные результаты могут быть получены при использовании Руководства Р 2.2.755–99 Госсанэпиднадзора Минздрава России.

В соответствии с данным Руководством можно оценить условия труда по 67 параметрам, 22 из которых составляют основу методики *количественной оценки напряжённости трудового процесса*. Каждый из 22 показателей оценивается одним из четырёх вариантов оценки (класс 1.0, 2.0, 3.1 и 3.2). Класс характеризует степень отклонения показателя от гигиенического норматива. Чем выше класс, тем выше степень отклонения.

Класс 1.0 – *оптимальные* условия труда, при которых сохраняется здоровье работников и создаются предпосылки для поддержания высокого уровня работоспособности.

Класс 2.0 – *допустимые* условия труда характеризуются уровнями факторов среды и трудового процесса, которые не превышают установленных гигиенических нормативов для рабочих мест. Возможные изменения функционального состояния организма восстанавливаются во время регламентированного отдыха или к началу следующей смены и не должны оказывать неблагоприятного действия в ближайшем и отдалённом периоде на состояние здоровья работающих и их потомство.

1 степень 3 класса (3.1) – условия труда характеризуются такими отклонениями уровней вредных факторов от гигиенических нормативов, которые вызывают функциональные изменения, восстанавливающиеся, как правило, при более длительном (чем к началу следующей смены) прерывании контакта с вредными факторами, и увеличивают риск повреждения здоровья.

2 степень 3 класса (3.2) – уровни вредных факторов, вызывающие стойкие функциональные изменения, приводящие в большинстве случаев к увеличению производственно обусловленной, появлению начальных признаков профессиональных заболеваний, возникающих после продолжительной экспозиции (часто после 15 и более лет);

3 степень 3 класса (3.3) – условия труда, характеризующиеся уровнями вредных факторов, воздействие которых приводит к развитию профессиональных болезней и потере профессиональной трудоспособности, росту хронической патологии.

4. Особенности умственного труда в условиях информационного и психоэмоционального напряжения

Метод гигиенической оценки и классификации условий труда по Руководству Р 2.2.755–99 позволяет оценить наиболее полный перечень факторов как в условиях изолированного воздействия, так и в составе комбинаций и сочетаний факторов. Недостаток этого метода связан с растущим объёмом подлежащих обработке данных для расчёта оценки условий труда. Данный недостаток может быть устранён разработкой специализированного программного обеспечения.

Однако основной причиной, препятствующей использованию метода в полной мере, является отсутствие возможности оценки индивидуального профессионального риска умственного труда, адекватного современным информационным перегрузкам. Для такой оценки необходимо определить, как с течением времени будет изменяться в заданных условиях труда состояние здоровья человека с заданной индивидуальной чувствительностью к факторам риска. Значимость этого вопроса связана с обеспечением безопасности человека в процессе умственной деятельности и предупрежде-

нием его о степени опасности выполняемой работы, о существующем риске.

У современных работников интеллектуального труда происходит тесное переплетение интеллектуальной и эмоциональной деятельности, что связано с достижением поставленной цели, преодолением затруднительных ситуаций и с социально-экономическими факторами риска. Нельзя полностью освободить работника от психоэмоционального напряжения. Однако необходимо повысить устойчивость адаптационных механизмов организма к эмоциональным стрессам упорядочить его деятельность. В современных условиях происходит постоянное усложнение профессиональной деятельности людей многих специальностей. Адаптационно-компенсаторные механизмы людей не справляются с повышенной информационной нагрузкой, что вызвано изменениями трудовых функций в связи с информационными процессами различной природы:

- восприятие и переработка разнообразной и многочисленной информации, которая часто имеет эмоциональную окраску;
- частое и быстрое переключение внимания, постоянная адаптация, быстрая реакция и выдача распоряжений подчинённым;
- острый дефицит времени для переработки значимой информации и для ответственных распоряжений;
- повышенная ответственность за принимаемые решения, особенно в экстремальных ситуациях, и частое возникновение состояния эмоционального стресса;
- необходимость длительного поддержания высокой интенсивности и напряжённости внимания, памяти, мышления, эмоции и различных функций организма;
- возникновение проблемной ситуации в условиях лимита времени и избытка информации;
- возникновение мотивационно-эмоционального поведения и образование доминанты для реализации интеллектуальных и эмоциональных актов высшего порядка;
- выполнение значительной части работы в ночное время;
- гипокинезия в связи с малым участием двигательного анализатора, поскольку работа выполняется большей частью сидя и не требует больших физических усилий;
- увеличение масштабов и усложнение содержания и структуры информационных потоков и всей информационной среды;
- взаимодействие психики человека с информационной средой отличается качественной спецификой и не имеет адекватных аналогов в информационном взаимодействии других биологических структур, технических, социальных и социо-технических систем;
- использование новых информационных технологий усиливает возможности информационно-психологического воздействия на людей. Различные способы манипулирования информацией реализуют *технология рефлексивного управления* для достижения

целей за счёт ущемления интересов отдельного человека или нескольких лиц.

Информационно-психологическую безопасность личности можно рассматривать как состояние защищённости её психики от действия многообразных информационных факторов, препятствующих или затрудняющих формирование и функционирование адекватной информационно-ориентировочной основы социального поведения человека, а также адекватной системы его субъективных отношений к окружающему миру и самому себе.

Информационная среда приобретает для человека характер второй, субъективной, реальности. Ту её часть, которая содержит информацию, неадекватно отражающую окружающий мир, и те её характеристики и процессы, которые затрудняют или препятствуют адекватности восприятия и понимания человеком окружающего мира и самого себя, можно условно обозначить как «иллюзорная реальность».

Несмотря на свою иллюзорность, и даже в силу своей иллюзорности, но в форме кажущейся реальности, она является основным внешним источником угроз информационно-психологической безопасности личности. Таким образом, в информационной среде в причудливых сочетаниях и многообразных формах переплетены адекватная и иллюзорная субъективные реальности. Они взаимодействуют и развиваются по собственным законам, создавая угрозы оставаться человеку полноценной личностью.

Источники угроз информационно-психологической безопасности личности заложены в самой биосоциальной природе психики человека, в особенностях её формирования и функционирования, в индивидуально-личностных характеристиках индивида. В силу этих особенностей люди отличаются степенью восприимчивости к различным информационным воздействиям, возможностями анализа и оценки поступающей информации и т.д.

Кроме индивидуальных особенностей, есть и определённые общие характеристики и закономерности функционирования психики, которые влияют на степень подверженности информационно-психологическому воздействию и присущи большинству людей. Существуют определённые закономерности восприятия и реагирования на малоосознаваемые и неосознаваемые воздействия, например на подпороговые стимулы и т.п.

Знание своих индивидуально-психологических особенностей и общих характеристик и закономерностей функционирования психики становится для человека в настоящее время не просто обязательным элементом его общей культуры, но и необходимым условием безопасности в социальном взаимодействии, в различных межличностных коммуникативных ситуациях.

Заключение

Одним из важных способов обеспечения безопасности человека в процессе трудовой деятельности является предупреждение и информиро-

вание о степени опасности выполняемой работы, о существующем риске повреждения либо утраты здоровья. Основной причиной, препятствующей развитию в полной мере концепции индивидуального профессионального риска, является отсутствие адекватных современных методов оценки риска. Современные профессиональные стандарты наиболее востребованных профессий в области информационных технологий однозначно определяют «специфические требования к состоянию здоровья деятельности работников»: *особых требований нет*.

Проблема информационно-психологической безопасности личности, её психологической защищённости и способов формирования психологической защиты в условиях информационного общества становится особенно актуальной как в теоретическом, так и в прикладном плане. Масштабность и мощность воздействия информационных факторов на психику людей выдвигает обеспечение информационно-психологической безопасности в современных условиях на уровень общенациональной проблемы.

Именно человек как личность и активный социальный субъект, его психика подвержены непосредственному действию информационных факторов, которые, трансформируясь через его поведение, действия или бездействие, оказывают дисфункциональное влияние на социальные субъекты разного уровня общности, различной системно-структурной и функциональной организации.

Общим источником внешних угроз информационно-психологической безопасности личности является та часть информационной среды общества, которая в силу различных причин неадекватно отражает окружающий человека мир. Выделение данного вида безопасности в качестве самостоятельного предмета теории и социальной практики связано с тем, что процессы и технология воздействия информационной среды на духовную сферу обладает качественной спецификой, которая определяет необходимость рассмотрения этой проблематики в концептуальном, методологическом и методическом планах.

УДК 378.6:004.78

В.И. Божич, М.Б. Савченко

ИНТЕЛЛЕКТУАЛЬНЫЕ ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА МЕНЕДЖМЕНТА ИНФОРМАЦИОННЫХ ПРОЦЕССОВ КОМПЬЮТЕРНОЙ ПЕДАГОГИКИ

Рассматриваются отличительные особенности компьютерного обучения, основанного на интеллектуальных инструментальных средствах.

Предлагаются принципы компьютерной педагогики с учётом психологических аспектов.

Введение. Научно-технические достижения создали *новый мир электронных средств доступа к культурным достижениям* и изменили привилегированные позиции книжной индустрии. Данное обстоятельство привело к необходимости пересмотра ключевых позиций образовательной деятельности на принципах книжной продукции. Созданная веками цельная система образовательной деятельности изменяется в сторону гибкой стратегии процесса обучения, в основе которой лежат информационные технологии.

Новая образовательная система является сложной диалоговой системой, основой которой являются концепции когнитивной психологии и систем поддержки принятия решений, распознавания образов и др. Новая система сохраняет процессы мышления на основе классического образования посредством сопоставимости слов устного и письменного выражения знаний. Однако многомодальные формы выражения знаний расширяются посредством *рефлексивных размышлений человека на уровне предлингвистических форм восприятия и понимания*. Такие формы могут быть выражены не только через слова и символы, а также через изображения (образы), звуки и другие особенности ассоциаций и действий.

Новая образовательная система изменяет практику ранжирования значимости студентов, построенную на принципах мотивации конкурентоспособности среди студентов. С точки зрения информационного управления, такая практика соревнования между студентами долгие годы считалась очень эффективной. Однако в последние два десятилетия возникла задача интенсивного поиска иных мотивационных источников для поощрения действий студентов не через соревнование, а через сотрудничество (совместную деятельность).

Причиной этому является система функционирования не образовательных организаций, а промышленных предприятий, административных учреждений и различного рода предприятий бизнеса, в основе которых заложен принцип сотрудничества, а не конкурентоспособности работников.

Сложные технологии привели к распределённым системам управления, диагностики и контроля информационных процессов, что, в свою очередь, потребовало от работников навыков совместной деятельности. Оценка деятельности работника определяется неконкурентоспособными испытаниями, характерными для образовательных систем. Она определяется его деятельностью в условиях выполнения работы совместно с другими людьми.

В настоящее время становится социально важной система преподавания, рассчитанная на развитие навыков сотрудничества, принципов

мотивации и оценки возможностей обучаемых с точки зрения совместной деятельности. Данное утверждение распространяется как на студентов, так и на педагогов.

1. Отличительные особенности компьютерной педагогики

Классическая система образования на основе печатной продукции состояла в системе авторитетного выбора учебного материала преподавателями и его передачи студентам с их ответственностью за его изучение. Такая система предоставляла каждому студенту намного больше материала, чем он может изучить. Система образования с правами выбора и изучения дополнительной учебной нагрузки предназначалась для одарённых студентов. Учебное заведение считалось лучшим в зависимости от дополнительного материала, который предоставлялся студенту для изучения. Выпускник такого учебного заведения считался рангом выше в сравнении с выпускниками других заведений.

Компьютерная система обучения также предоставляет студенту большое количество учебного материала, причём гораздо большее, чем любой индивидуум может обработать (изучить), и не исключает потребность вдохновения и усердия в работе студента и в деятельности преподавателя. Однако процесс такого образования имеет характерные отличительные особенности:

- каждый студент не только изучает учебный материал, но и развивает уникальные, присущие только ему, способности выбора соответствующего материала;
- задача педагогической системы не может и не должна определять точно, что студент выбирает, но обязана помочь ему извлечь максимально возможную пользу при его самообучении на основе элементов «конструктора знаний», которые выбраны студентом.

При компьютерном обучении возможности авторитетного (индивидуального) выбора студентом материала резко возрастают. Студент перестаёт быть послушным учеником. Он больше не ответствен только за изучение учебного материала, его ответственность распространяется на уникальный и полезный выбор материала с целью его разумного исследования.

Таким образом, процедура обучения становится в ранг экспериментального исследования.

Изучение и организация учебного процесса строятся на основе трёх ключевых действий:

- постановка вопроса, которая определяет запрос учебного материала;
- выбор, оценка и изучение материала с помощью тех или иных способов;
- представление результатов обучения способами, которые могут быть оценены различными средствами.

Роль преподавателя значимо изменяется. Преподаватель становится «путеводителем», который несёт студенту интерпретацию его авторского замысла на процесс обучения через побуждение интереса и значимости учебного материала, через объяснение целей и задач, предостережений об ошибках и др. Указанная роль преподавателя расширяется проверкой знаний, навыков и умений студентов.

Информационные технологии компьютерного образования требуют специальной организации структуры данных, а также наличия интеллектуальных инструментальных средств, отражающих авторский замысел и стратегии преподавателя, а также желаемые процедуры действий студента. Более того, они должны не только предоставить знания, но и выявить их понимание.

Большинство форм вычисления, корреляции, комбинации и подключения, которые могут делать люди, компьютеры могут делать лучше, расширяя способности человека сортировать, заказывать, ранжировать и выбирать информацию. Компьютерное обучение ориентировано на интеллектуальные инструментальные средства, отличительными характеристиками которых является требование *нелинейного процесса обучения*. Такие средства должны решать задачи обслуживания индивидуальных запросов студента на информацию относительно цельной системы знаний, хранимых в различных внешних структурах. Важно заметить, что функции инструментальных средств компьютерного обучения должны восприниматься студентами как инструмент, а не являться дополнительным приложением для изучения.

Перечисленные выше особенности указывают на существенные отличительные черты компьютерного образования от классической парадигмы системы образования. В такой ситуации известные рабочие подходы классической педагогики могут быть использованы только частично и очень частично. Дальнейшие исследования и результаты в научных областях искусственного интеллекта и когнитивной психологии могут служить основанием практического создания новой системы образования студентов на базе электронной и информационной продукции в отличие от классической системы обучения, построенной на базе книжной полиграфической продукции.

2. Принципы компьютерной педагогики

Система компьютерного обучения трактуется не просто как оппозиция, альтернатива классической системе образования, а как переворот системы образования и культуры в целом. Условно можно выделить следующие принципы новой стратегии образования:

2.1. Индивидуальное обучение. Имеет место двоякость понимания данного принципа. С одной стороны, имеет место реальная возможность использования образовательного процесса в направлении обучения одного

индивидуума (или небольшого числа студентов), в сравнении с ранними образовательными стандартами – классная комната рассчитывалась на 24 учащихся. С другой стороны, индивидуальность обучения рассчитана на активную позицию самого обучаемого индивидуума в рамках новой образовательной деятельности – усилилась значимость понятия «*самостоятельная работа обучаемого*».

2.2. Конструктивно-инструктивное представление процесса образования. Используемая ранее командная система обучения учащихся в классной комнате изменяется на представление преподавателем учебного материала в виде «*конструктора для сборки знаний*» самим учащимся. Авторский замысел преподавателя представляется инструкциями к пользованию «конструктором знаний» в рамках той или иной предметной области. Сказанное означает, что профессия преподавателя в новой системе «электронного» образования требует пересмотра классических, устоявшихся требований к профессиональной деятельности преподавателя. В новой образовательной системе студент обладает *правом самостоятельного принятия решений* при работе с «конструктором знаний». При этом его творческий уровень и подход может изменить, а порой даже исключить авторский замысел преподавателя – систему усвоения знаний по заданной предметной области.

2.3. Представление предметной области с различными уровнями её детализации. Релевантность уровня изучаемой предметной области определяется как образовательной системой, так и самим обучаемым. Другими словами, новая педагогическая система отменяет принцип *последовательного изучения предметной области*, представленной минимальными объёмами информации (текст, подраздел, раздел, набор тем и т.д.).

2.4. Индексная градуировка (адресация) предметной области. Поиск – фундаментальная проблема всех культур. При цифровом способе кодирования информации данная проблема является особенно существенной, так как для управления поиском информации необходимо иметь мощные инструментальные средства в связи с решением задач выборочного её получения и обмена информацией. Исторически поиск информации развивался относительно её текстовой формы представления. Текст использовался и до сих пор используется, чтобы каталогизировать большинство других форм – карты, изображения, числовые таблицы, фильмы, регистрация и др. Однако с развитием информационных систем оказалось, что такая форма оказалась не единственной формой интеллектуального выбора и поиска.

В последнее время стали применяться методы *нетекстового поиска*, основанные на свойствах когнитивного (образного) восприятия информации путём визуализации данных пространственной памятью человека через символы, звуки, геометрические отношения и др.

Когнитивные методы поиска оказались эффективными для компьютерных обучающихся систем в сравнении с устным или письменным формированием направления поиска.

Таким образом, при использовании компьютерных (машинных) систем обучения *сортировка учебного материала* для изучения производится не на основе декларативных знаний, что имело место в книжной продукции, а на основе различных параметров – индексов (значки, изображения, звуки, анимационные жесты и др.), через которые представляются знания.

Следует заметить, что такие индексы являются новым элементом современной «электронной культуры» и определяют процедурные знания для поиска «истинных» декларативных знаний в виртуальном мире знаний. Указанные индексы формируют новый словарный запас современного учащегося.

2.5. Интеллектуальная синтетическая организация знаний. Самостоятельная работа обучаемого с «конструктором знаний» приводит к результату *интегрированного образования*. Интеллектуальный синтез знаний осуществляется собственными усилиями учащегося.

Другими словами, интегральное образование приводит к «объединённой индивидуальности» учащегося, выраженной набором навыков, умений и знаний. Указанный набор в целом характеризует индивидуальную особенность индивидуума, объединяющую в себе набор предпочтений относительно общечеловеческой культуры.

2.6. Интеграции процесса образования путём формирования внимания. Интегрированное образование является активным образованием со стороны студента. Процесс его приобретения является достаточно сложным в сравнении с классическим процессом образования, так как преподаватель строит «конструктор знаний» в ориентации на многостороннюю, разнообразную деятельность студента, но при этом не объединяет весь учебный материал.

Психологи и педагоги отмечают три возможные формы интеграции процесса образования: а) комбинаторная интеграция; б) саморефлексивная интеграция; в) реинтеграция.

Комбинаторная интеграция начинается с детства и продолжается в течение жизни человека. Она заключается в формировании новых навыков человека относительно его «объединённой индивидуальности» через *фокусировку внимания на событиях и фактах*.

Саморефлексивная интеграция начинается в юности и связана с сознательным формированием профессии человека. Такая интеграция отражает интересы и утверждения индивидуальности на уровне его сознания. Она строится на практическом опыте человека и осуществляет проективное воображение, видя различные возможности, разворачивающиеся в

обозримом будущем. В отличие от простых комбинаторных суждений саморефлексивная интеграция связана с процессом выбора и принятия решений.

Необходимость выбора решения приводит к *формированию внимания на потребностях в критериях и принципах, требует определения целей на сознательном уровне.*

Саморефлексивная интеграция сводится к формированию человеком синтетической модели мира и определяет критерии и принципы управления этой моделью – знания, умения, навыки.

Реинтеграция связана с процессом трансформации синтетической модели путём комбинирования отдельных подмоделей под действием новой целеустремлённой ориентации.

Такое преобразование модели является следствием *формирования внимания на результатах опыта, полученного не в результате управления и активной позиции человека*, а в результате болезни либо экстремальных изменений обстоятельств, которые резко разрушают синтетическую модель навыков, знаний и умений. Другими словами, реинтеграция связана с трансформацией оценки внешних вызовов на подсознательном и сознательном уровнях.

2.7. Психологическая безопасность процесса образования. Создание средств гипер- и мультимедиа привело к представлению нашей культуры в цифровом коде. Дальнейшее развитие таких систем приводит к расширению культурных возможностей. Действительно, компьютер и системы мультимедиа позволяют упростить и расширить интеллектуальную деятельность человека, а также сделать беспрецедентным доступ людей к культурным ценностям.

Смысл и назначение системы гипермультимедиа – это практическое объединение в одно целое всех форм информации. При этом производство и воспроизводство информации не является только техническим процессом. Это процесс, управляемый стимулами личностей в соответствии с некоторыми законами.

Компьютерные информационные технологии не только позволяют ускорить интеллектуальные достижения человечества, но и усиливают ответственность людей за их поведение.

В этой связи высокую значимость приобретает деятельность образовательной системы. Поскольку влияние информационных технологий расширяет интеллектуальные способности, следствием этого является необходимость перестройки системы образования в направлениях техники обучения и усиления ответственности за знания и технологические возможности их использования.

Традиции классической системы образования строились на объединении целостной системы знаний, передача которых осуществлялась через общение учащихся с педагогами. Интегральное образование резко отличается от таких педагогических приёмов и ставит обучающегося в условия психологического напряжения, предоставляя студенту самостоятельно получать знания через компьютерную систему обучения с помощью «конструктора знаний».

Эффектом от такого психологического напряжения является возможность учащегося *самостоятельно синтезировать своё образование* в зависимости от его индивидуальности, выраженной синтетической моделью его мира (навыки, цели, вера и др.).

Возможность самостоятельной работы высоко оценивается интеллектуальной личностью. По результатам социологических исследований замечено, что в настоящее время большинство людей на первое место выдвигают образовательную возможность вместо социального обеспечения, карьеры, материальных выгод и других ключевых понятий, на которых ещё недавно строились принципы развития цивилизации. Последнее отражает новое явление и, несомненно, связано с *понятием психологической безопасности личности и общества*.

3. Психологические аспекты компьютерного обучения

В течение длительного периода психологические исследования позволили обосновать и создать педагогическую систему организации управления учебной деятельностью учащихся посредством классных комнат. Однако информационное общество существенно отличается от ранее созданной цивилизации новым миром электронной культурой. Возникло явление, которое получило название «*информационный взрыв*».

Объём всей информации, накопленной человечеством, делается всё менее доступным отдельному индивиду. Этот процесс совершается не в арифметической, а в геометрической прогрессии.

Современный городской житель на протяжении недели получает и обрабатывает в своём мозгу столько же информации, сколько житель XVI или XVII вв. за всю свою жизнь.

Аналитические прогнозы утверждают, что именно инновации в информационных и коммуникационные технологии, рост уровня и мощности телекоммуникационных сетей будут играть ведущую роль в изменениях, происходящих в обществе и в жизни каждого человека.

По указанным причинам психологические исследования новой системы компьютерного образования являются многопрофильными и опираются на результаты научных исследований технических и гуманитарных областей: физиология и клиника развития мозга человека, организация общества, социология, антропология, философия, информационное моделирование сложных систем и др.

Целью проводимых системных исследований является разработка принципов обучения, позволяющих сохранить «умственное здоровье» уча-

щихся и организовать эффективное обучение различных групп учащихся с точки зрения способностей к усвоению материала, к информационным перегрузкам и др.

Значимость психологических исследований определяется тем фактором, что новые компьютерные образовательные системы сосредоточены на человеческих функциях, которые должны быть учтены на различных уровнях функционирования обучающейся системы: управление обучением, изучение, оценка знаний. Менеджер информационных процессов образовательной системы должен учитывать психологические принципы для организации компьютерного обучения, которые должны быть ориентированы на внутренний мир студента в условиях его взаимодействия с внешней средой. Поэтому новая система образования должна преследовать цель реализации психологических принципов в педагогической практике с учётом значимых факторов:

- познавательные факторы;
- эмоциональные факторы;
- факторы развития;
- личностные и социальные факторы;
- индивидуальные различия студентов.

Указанные факторы позволяют вести разработку «интеллектуальных электронных учителей» образовательной системы с учётом «человеческого фактора». Данное обстоятельство указывает на необходимость разработки тезауруса декларативных и процедурных знаний, которыми должны обладать модели тьюторов с целью управления деятельностью студента:

- вовлекать студентов в учебный процесс с учётом их самостоятельной учебной деятельности и проявления индивидуальных наклонностей;
- поощрять студентов в осуществлении взаимосвязи полученной новой информации с имеющимися знаниями;
- обеспечивать стимулирование и руководство деятельностью студентов на уровне мышления в рамках укрупнённых концепций с целью приобретения и интеграции знаний;
- включать конструктивную и информативную обратную связь на уровне развития практических навыков принятия решений;
- предлагать возможности для приобретения и реализации различных стратегий и альтернатив изучения материала, относящегося к различным областям знаний;
- поощрять решение проблемы, планирование, сложное принятие решения, обсуждение проблемы с другими людьми с целью развития способностей совместной деятельности;
- помогать студентам в развитии психологических навыков с целью понимания и уважения индивидуальности других людей;

- позволять студенту планировать изменения учебного плана специальности, в рамках которой он проходит обучение с целью повышения эффективности процесса самообучения;
- гарантировать, что деятельность студента будет оценена правильно, что его индивидуальность будет наблюдаться образовательной системой в целом и менеджерами его обучения (понятие более широкое, чем преподаватель), что студенту будет предоставлена возможность активного моделирования его будущей профессиональной деятельности с целью развития его практических знаний, умений и навыков.

Выполнение указанных выше требований связано с комплексом мероприятий, однако следует выделить наиболее значимое мероприятие с точки зрения психологии – эффективная оценка знаний студентов. Такая постановка значимости «оценки знаний» является естественной для обучающей системы, центральным звеном которой является физическое лицо – студент. Механизм оценки знаний является тонким механизмом психологического воздействия на студента.

Форма и содержание понятия «*оценка знаний*» является точкой отчёта деятельности студента и позволяет стимулировать и управлять дальнейшим процессом его обучения. Важным фактором такого механизма является вовлечение студента непосредственно в процесс оценки их деятельности с учётом их индивидуальных целей обучения и индивидуальных успехов, а не оценка знаний относительно стандартов и некоторых нормативных методов аттестации.

Анализ результатов психологических исследований по влиянию оценки знаний на деятельность студентов для образовательных компьютерных систем обучения позволяет сформулировать следующие принципы, которые должны быть использованы при построении компьютерных моделей оценки образовательной деятельности студента:

- основная цель механизма оценки деятельности студентов должна быть ориентирована на дальнейшее стимулирование студента в получении знаний;
- оценка знаний студента должна соответствовать признакам мотивации учебной деятельности студента, которые определяются списком некоторых стандартов целесообразности и превосходства с точки зрения жизненных целей и устремлений индивидуума. Такие стандарты складываются из множества требований к деятельности студента на уровнях государства, вуза, преподавателей, родителей и др. Важным обстоятельством при этом является чёткое определение стандартов превосходства и гарантия их выполнения;
- оценка должна обеспечить доверие и обоснованность относительно широкого спектра достижений студента в рамках учебного плана;

- оценки должны измерить побуждения студента, его эмоциональные реакции так же, как и его познавательные знания, навыки и умения;
- оценки должны включить условия для многократных альтернативных ответов и возможности творческого роста студента на основе пояснения его «ошибок» с целью дальнейшей его самостоятельной учебной деятельности;
- оценка должна выявлять подлинное усилие студентов, их побуждение и обязательства к целям обучения, а также способствовать самооценке и процессу самостоятельной работы (обучения);
- результаты оценки должны быть ясными для студента. Оцениваемые знания, умения и навыки студента должны соответствовать ежедневному учебному графику его деятельности;
- оценки должны обеспечить ежедневную обратную связь между студентом и обучающей системой;
- оценки деятельности студента должны быть периодически представлены для различной категории потребителей выпускников вуза: руководители вуза, преподаватели, родители и заинтересованные предприятия и физические лица.

Заключение

Разработка и внедрение новой педагогической системы компьютерного обучения ведётся в направлении кардинального изменения и замены старой системы организации идей и знаний, отменяя привычные принципы классического образования с помощью классных комнат, расписания, уроков и др.

Информационные технологии и асинхронная связь улучшают способности людей работать совместно и интерпретировать взаимные результаты.

Актуальным становится понятие «менеджер информационных процессов образовательной системы». Реальность функционирования компьютерной системы обучения связана с двумя факторами деятельности менеджеров:

- формирование внимания студента на образовательном процессе не возникает само по себе и требует командного подхода;
- результат обучения всегда связан с категорированием студентов по мотивации и результатам обучения и требует наличия обратной связи с целью оценки и направленности деятельности студента.

Для создания и развития компьютерной образовательной системы необходима разработка двух взаимосвязанных моделей: модель преподавателя и модель студента. Достоверность таких моделей позволит целенаправленно вести процесс образования.

Возможности современных информационных технологий мультимедиа, нейрокомпьютинга, эволюционного поиска и генетической оптимиза-

ции, мультиагентных и бионических систем указывают о реальности создания естественного интерфейса «студент-преподаватель» и интеллектуальных инструментальных средств компьютерных образовательных систем.

УДК 338.24:004

З.И. Попова, М.Б. Савченко, В.И. Божич

ЭМЕРДЖЕНТНЫЙ ИНТЕЛЛЕКТ ИНФОРМАЦИОННОГО МЕНЕДЖМЕНТА

Предлагается вариант построения доски объявления мультиагентной системы, учитывающей симбиоз образного и символично-логического представления знаний и их обработки.

Введение. Эмерджентный интеллект – это явление, определяющее процесс самоорганизации, возникающий за счёт множества скрытых или явных условий, сложившихся в ситуации, спонтанно и в заранее непредвиденный момент времени. Особенностью эмерджентного интеллекта является непредсказуемость и динамика процесса принятия решений посредством большого количества взаимодействий (сотен и тысяч), которые почти невозможно отследить.

Понятие эмерджентного интеллекта связывают с мультиагентными системами [1, 2], которые могут обеспечить интеллектуальный информационный менеджмент при решении задач поиска, сбора информации и фильтрации данных, особенно в виртуальном пространстве Интернет. Элементом такого интеллекта является «социальный агент», который взаимодействует с другими агентами – программными роботами или людьми. Необходимо учитывать, что взаимодействия социального агента ориентированы на достижения целей более крупной мультиагентной системы – социальной системы (совокупность социальных явлений и процессов, которые находятся в отношениях и связи между собой и образуют некоторый целостный социальный объект).

Развитие информационного общества усложняет функции информационного менеджмента и задачи, решаемые эмерджентным интеллектом на основе мультиагентных систем, по следующему ряду причин:

- множество значений для большинства показателей является нечётким;
- отсутствует чёткое значение удельного веса этих показателей в совокупной оценке эффективности деятельности;
- значимость показателей является переменным параметром, т.е. определяется неявным образом и зависит от параметра времени;

- статистические данные являются ограниченными и известны лишь для небольшого количества возможных видов деятельности;
- взаимосвязи диапазонов показателей и соответствующие им оценки эффективности деятельности известны только в виде экспертных правил, как правило, в лингвистической нечёткой форме;
- экспертные отношения предпочтения на множестве показателей субъективны в некоторых диапазонах их значений.

Агенты функционируют в виртуальном мире, решая заданную проблему, и взаимодействуют между собой путём посылки сообщений. Агентам определяют цели, которые они должны достигать, но не определяют сценарии исполнения задач по достижению этих целей. Эти сценарии исполнения задач формируются и исполняются агентами самостоятельно. Агенты могут соревноваться друг с другом или сотрудничать в процессе принятия решения. Как результат переговоров агентов, формируется текущее решение проблемы, которое гибко меняется в соответствии с динамикой среды.

На каждом шаге агенты рассматривают входы системы и реагируют на непредсказуемые события. Реакция может быть самостоятельной или осуществляться во взаимодействии с оператором – лицом, принимающим решение (ЛПР). Влияние и значимость человеческого фактора на процесс принятия решений накладывает отпечаток на развитие эмерджентного интеллекта, принципиальным моментом которого является снятие противопоставления между «исследователем» и «объектом исследования».

1. Агентно-ориентированная парадигма

Современные результаты когнитивной психологии и искусственного интеллекта позволяют ставить и решать новые задачи построения систем, наделённые психикой деятельности. При этом важным вопросом является выбор интеллектуальных информационных средств, которые будут удовлетворять целостности объектов, порождаемых естественным и искусственным интеллектами.

Возможные пути решения задач по выделению указанных отношений должны быть основаны на симбиозе образного и символично-логического представления знаний и их обработки в системах искусственного интеллекта (ИИ). В этой связи разработка новых технологий принятия решений должна учитывать результаты агентно-ориентированного подхода, который позволяет учесть представление ментальных и интенци-

ональных моделей деятельности ЛПР и системы ИИ в форме нечётких понятий и правил лингвистического и когнитивного характера.

Использование результатов такого подхода позволит учесть значимость знаковых представлений семиотики с точки зрения формирования взаимодействия образного мышления ЛПР и символьного мышления системы ИИ. При этом аспект семиотического рассмотрения образа позволит трансформировать проблему значимости образа в проблему координации деятельности. Такая оценка значимости образа позволит использовать концептуальные модели в виде наглядного представления семантических отношений «совместной деятельности» на уровнях образного и символического мышлений «агент – действие – объект».

Концептуальная модель, отражающая агентно-ориентированную парадигму, показана на рисунке.

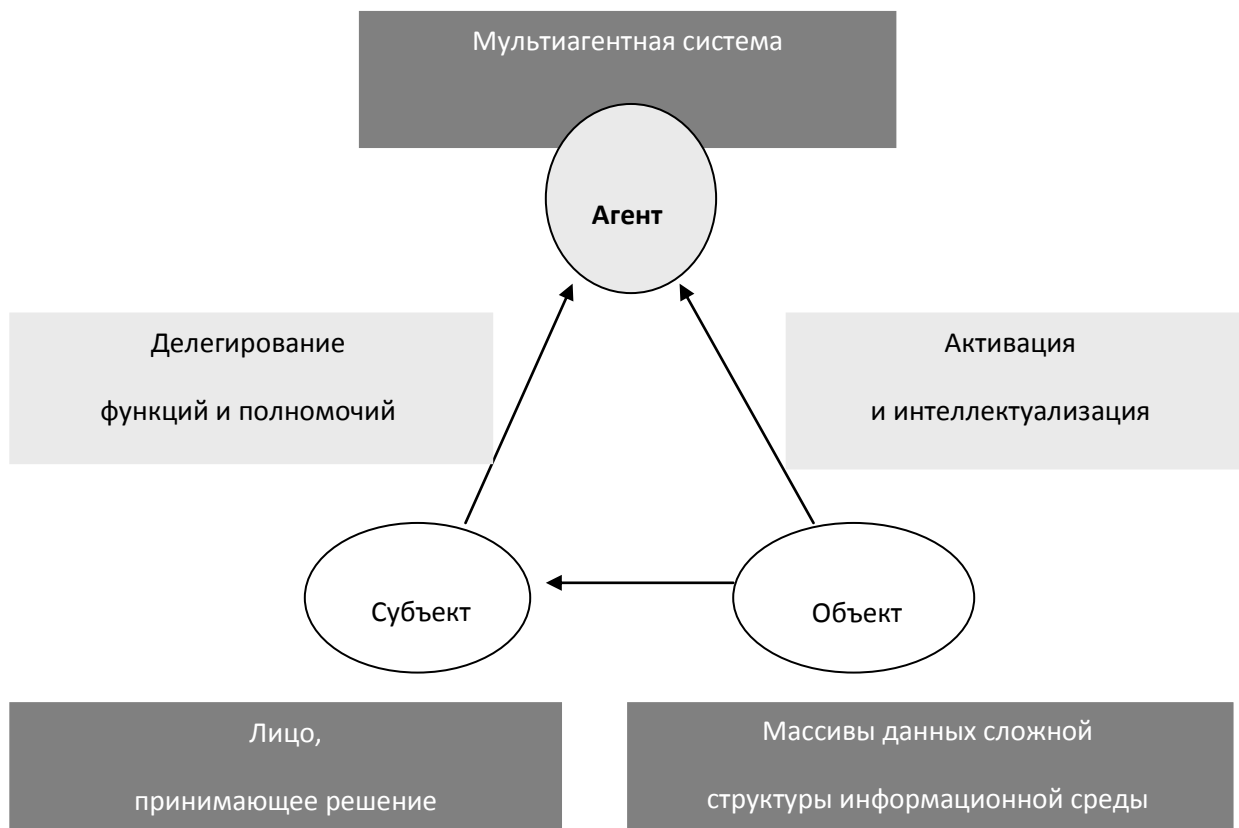


Рис. Концептуальная модель агентно-ориентированного подхода

В данной модели «агент» является значимой компонентой, так как он является промежуточным звеном между субъектом и объектом. В такой модели агент взаимодействует на основе «договорных отношений» между субъектом и объектом и стремится синтезировать «искусственного экс-

перта», способного самостоятельно решить задачу фильтрации полезной информации, соответствующей тематическому запросу.

С позиции субъекта агент представляет собой «квазисубъект» на уровне образного мышления, который имеет определённые обязательства перед ЛПР, и действующий по его поручению. С позиции объекта агент является «истинным объектом» на уровне символьного мышления, способным формировать действия для достижения определённых целей.

При таком подходе мультиагентная система должна удовлетворять целостности объектов, порождаемых естественным интеллектом, и моделировать процедуры управленческой деятельности, наделённые психикой деятельности человека посредством множества пар «цель – действие». Понятие «агент» определяет единение между субъектом (ЛПР) и объектом (системой ИИ) и является открытой системой, способной воспринимать информацию с ограниченным разрешением, обрабатывать её, преследовать свои собственные цели и действия.

2. Отличительные особенности подсистем хранения и обработки информации человеком и компьютерной техникой

Важно отметить, что системы хранения и обработки знаний человека и компьютерной техники базируются на двух различных подсистемах. Для человека образное мышление является врождённым, а поэтому первичным. Символьное мышление возникает в процессе обучения человека через взаимосвязь с символической средой, представленной текстами, речью и др. Для компьютерной техники имеет место противоположная подчинённость подсистем. Первичной является символическая подсистема, а образная подсистема (изображения, графика и др.) является вторичной и базируется на символических конструкциях и методах. Результатом такой противоположности является факт различной сложности выполнения операций хранения и обработки информации человеком и компьютерной техникой.

Построение новых информационных технологий за счёт слияния естественного и искусственного интеллектов, несомненно, должно считаться с проблемой противоположности значимости подсистем хранения и обработки информации человеком и компьютерной техникой. В связи с этим важно учитывать результаты когнитивной психологии (экспериментальные исследования, информационные модели и др.), которые дают некоторое основание и направления решения проблемы слияния образного и символического мышления [3]. При этом одной из центральных является проблема совмещения механизмов обработки образов человеком и компьютерной техникой через формирование внимания ЛПР. Решение данной проблемы возможно средствами моделирования «совместной деятельности» человека и компьютерного устройства в рамках «субъектно-ориентированной» парадигмы «совместной деятельности» ЛПР и системы ИИ [4].

Процесс восприятия информации человеком основан на механизме «формирования внимания» [5–7]. Результаты исследования такого механизма сводятся к некоторым экспериментальным фактам и информационным моделям восприятия информации человеком. Описание таких моделей и механизмов связано с особенностями работы мозга естественного интеллекта по интерпретации и распознаванию образов. При этом восприятие информации определяется как процесс её соотнесения к некоторому классу и наделения её некоторым смыслом, или, иначе, как процесс построения внутреннего представления образа:

- внутренний образ, хранящийся в памяти человека, является сугубо внутренним представлением и определяется динамично изменяющейся в процессе жизни системой категорий индивидуального человека;
- внутренний образ в памяти человека не может быть объективирован. По этой причине совместная деятельность современных людей не основана на обмене образами, а достигается способами представления информации через слова, текст, рисунок и т.п.;
- объективизация образа через символьную информацию фактически исключает понятие внутреннего образа и тем самым приводит к потере и искажению информации. Действительно, порой краткая образная мысль в мозге человека требует длительного и тщательного описания на языке символов, привычных для общения между людьми;
- процесс формирования цепочки «информация – внутренний образ – информация» не сводится к однозначному кодированию и декодированию информации через внутренний образ;
- мозг человека работает по семантическому принципу, определяя смысловое значение образа, и существенно отличается в средствах представления и обработки информации – образов, характерных для компьютерной техники. Понятие «информация» – образ как внутреннее состояние некоторого процесса мышления человека требует особых описаний;
- процесс «восприятия» информации в компьютерных технологиях сводится к преобразованию информации без её интерпретации на основе синтаксических манипуляций с символами без учёта их семантического смысла.

Из сказанного следует актуальность использования особенностей образного мышления человека при создании новых информационных технологий. Субъективность, целостность и высокая скорость интерпретации образа человеком в сравнении с символической обработкой информации в компьютерах должна быть учтена при организации эмерджентного интеллекта. При этом деятельность ЛПР необходимо свести к учёту когнитив-

ных свойств информации (свойств визуализации) как на этапе её восприятия, так и на этапе взаимодействия с системой ИИ через когнитивную информацию. Другими словами, ЛПР должно обладать возможностью работы с визуальной информацией при её восприятии и изменении, опираясь только на образное мышление, исключая целенаправленный перевод информации в символьную форму на этапе формирования связи «внутренний образ – информация».

В рамках такого подхода представленная когнитивная информация воспринимается со стороны ЛПР на уровне образного мышления, а со стороны системы ИИ – в виде символьной формы. При этом предполагается, что ЛПР имеет возможность внести некоторые изменения на рисунке (изображении), фокусируя внимание на рисунке и сохраняя тем самым образное мышление. Все действия ЛПР преобразуются интерфейсом в символьную форму для учёта соответствующей реакции системы ИИ на действия ЛПР. Также должна решаться и обратная задача – представление действий системы ИИ для ЛПР в виде изменённой когнитивной картинке, а не символьной информации в виде матриц, списков и т.п.

3. Монитор доски объявления эмерджентного интеллекта

По утверждениям психологов [6–9], визуализация представления информации является инструментом, позволяющим ЛПР ускорить и повысить эффективность решений. Акт формирования внимания является актом интерпретации образа и является рассуждением, обобщением, формированием гипотез. Использование образного мышления ЛПР при работе с системой ИИ даёт основание для построения новых информационных технологий на базе слияния естественного и искусственного интеллектов.

Для организации такой деятельности в системе ИИ должно быть введено представление агента – *«искусственного двойника ЛПР»* – для учёта и оценки деятельности *«естественного»* ЛПР, который должен решать вопросы соотнесения своих действий с нормативной базой, заложенной в основы функционирования системы ИИ. Постановка вопроса о разработке моделей автоматов, имеющих ментальные (психические) характеристики, рассматривалась ранее, например в работах [10–14]. Авторами этих работ были предложены описания поведенческих актов с помощью фреймов поступков, которые заданы специальными *графами «замысел – реализация»* с такими характеристиками, как мотивы, намерения, нормы, ожидаемые оценки, потребности и др.

В какой-то мере эти работы явились основой развития агентно-ориентированного подхода в рамках направления «искусственный интеллект» [15, 16]. Важным аспектом указанного подхода явился переход от рассмотрения управления с позиции автоматной модели (как последовательности актов выбора решений) к распределённой системе, в которой структуры управления являлись множеством «паттернов прохождения сообщений», совместно решающих коммуникационные проблемы путём об-

мена сообщениями в контексте семантики открытых систем. По этой причине появилось понятие «доски объявлений» [16], т.е. зоны совместной работы в базе данных на основе модели пространства гипотез и решений.

Монитор как блок архитектуры доски объявлений используется для контроля всего процесса решения задачи в распределённой системе ИИ; сортирует задачи из списка и, как правило, эвристическим путём выбирает одну из множества следующих задач для выполнения. Монитор осуществляет проверку доски объявлений и области сообщений после работы каждого агента, анализируя слоты классной доски, связанные с гипотезами, определяющими текущее состояние процесса решения в соответствии с некоторым планом – сценарием и таким образом определяет, какие модули знаний соответствуют текущему состоянию.

Программные агенты распределённой системы ИИ взаимодействуют на основе модели «договорных отношений» и стремятся все вместе синтезировать «искусственного специалиста», способного самостоятельно решить задачу. Действия, инициируемые монитором, постоянно меняются. В процессе решения задачи система ИИ всё время инициирует различное мнение агента – «искусственного эксперта» – и(или) коллективное мнение различных «экспертов», каждый из которых моделируется подобно фрейму множеством пар «атрибут – значение» и может обращаться за сведениями к другим специалистам.

Заметим, что процессы такой системы ИИ не отражают роль и значимость ЛПР, так как не отражают принципы «субъектно-ориентированной парадигмы» по причинам главенства идеологии нормированной системы ИИ. Однако имеется возможность модификации процессов распределённой мультиагентной системы с целью реализации принципов «субъектно-ориентированной парадигмы» посредством расширения функций монитора доски объявлений введением доступа к ней со стороны ЛПР.

Следует обратить внимание на схожесть функций и различие представления информации в мониторе доски объявлений и в мониторе ЛПР. Монитор доски объявлений по своему функциональному назначению рассчитан как средство «визуализации символьной информации» для агентов – средств обработки двоичной информации. Монитор для ЛПР следует рассматривать как прибор «визуализации графической информации». В целом, монитор доски объявления эмерджентного интеллекта может учитывать стратегии и тактики деятельности, определяемые как когнитивными свойствами ЛПР (на основе визуальной информации на мониторе), так и действиями агентов (благодаря символьной информации доски объявлений). Совместная деятельность агентов и ЛПР должна определяться некоторым протоколом коммуникации и учитывать психологические характеристики. Данное обстоятельство является важным с точки зрения формирования внимания ЛПР и учёта его когнитивных особенностей.

В связи с этим необходимо учитывать типологию агентов с позиции взаимодействия «субъект-объект». Уровень субъективности агента определяют либо только символическими представлениями, требующимися для организации рассуждений, либо, в противоположность этому, агент работает только на уровне образов. Такая особенность деятельности указывает на необходимость использования когнитивных агентов, имеющих символическую модель внешнего мира, на основе которой агент проводит собственные рассуждения, принимает решения или выполняет действия, изменяющие среду. Отличительной особенностью когнитивных агентов является наличие ментальных свойств, которые относятся к категории психических характеристик человека, – желания, намерения, цели, убеждения и др.

Использование когнитивных агентов предопределяет: а) наличие системы ценностей и вариантов поведения и б) коммуникацию ЛПР и агентов с учётом значимости знаковых представлений семиотики для взаимодействия образного мышления ЛПР и символического мышления системы ИИ. Моделирующая подсистема когнитивного агента должна отражать семантические конструкции «агент – действие – объект» – графическое изображение совокупности знаков и интенциональных отношений между ними, позволяющими связать динамические отношения соучастия в совместной деятельности ЛПР и системы ИИ. Генерацию возможных решений (ситуационное управление) можно реализовать с помощью когнитивной карты [17] – средства «фокусировки внимания» образного мышления ЛПР.

Как известно, когнитивная карта представляется в виде графа и позволяет описывать субъективное восприятие ЛПР (или группы ЛПР) сложного объекта, проблемы или функционирования некоторой системы. Такая карта предназначена для выявления структуры причинных связей между элементами сложного объекта и оценки последствий, происходящих под влиянием воздействия на эти элементы или изменения характера связей. Альтернативы, разнообразные причины, эффекты, цели и необходимые средства их достижения могут рассматриваться как переменные концептов. В этом случае на когнитивной карте они представляются вершинами. Их влияние друг на друга обозначается направленными дугами, отмеченными определёнными знаками, которые определяют характер влияния.

Существует несколько методов построения когнитивных карт: либо когнитивную карту строит сам ЛПР, либо когнитивная карта создаётся на основе изучения документов одним или несколькими экспертами. Преимущество второго метода заключается в возможности обобщать индивидуальные мнения и представлять более широкий диапазон оценок. Предложенная выше идея рассмотрения монитора доски объявления эмер-

джентного интеллекта ориентирована на совместную деятельность по разработке когнитивных карт множеством программных агентов и ЛПР.

Модель предметной области информационного менеджмента в виде когнитивной карты отражает параметры событий и степени их взаимного влияния как точными количественными параметрами, так и нечёткими качественными соотношениями. Это особенно важно при моделировании ситуаций, когда приходится оперировать такими зыбкими ментальными и интенциональными понятиями. Важно заметить, что практика использования «нечётких когнитивных схем» продемонстрировала, что они являются активным интеллектуальным средством оценки ситуации и генерации решений, основанных на эвристических предпочтениях ЛПР.

1 этап – формирование когнитивной карты. На этом этапе определяются основные факторы, влияющие на решение проблемы и их причинно-следственные связи.

2 этап – создание базы знаний экспертной системы. После того как на основе когнитивного анализа выработаны предложения или приняты решения о выполнении определённого набора действий, в базу данных экспертной системы записываются условия, при которых могут быть выполнены эти действия и детали их выполнения в соответствии с создавшимися условиями. На основе этой информации, записанной в базе знаний, экспертная система в соответствии с конкретной обстановкой генерирует решение о порядке выполнения операций.

3 этап – создание сценариев – мозаик операций. Система поддержки принятия решений формирует возможные сценарии – последовательности выполнения таких операций (действий). Сценарии могут различаться не только последовательностью действий, но и набором тех или иных операций.

Заключение

Разработка новых технологий информационного менеджмента должна учитывать концепцию агентно-ориентированной организации эмерджентного интеллекта. Для решения задач управленческой деятельности системная организация такого интеллекта должна отражать субъектно-ориентированный подход в ориентации на персональную поддержку деятельности ЛПР. Она должна содержать в себе интенциональные процедуры (наделённые психикой) и когнитивные процедуры деятельности ЛПР и когнитивных агентов.

Сочетание естественного и искусственного интеллектов в рамках одной системы эмерджентного интеллекта позволяет создать условия для развития профессиональных и личностных характеристик ЛПР и новых приёмов анализа и фильтрации полезной информации в массивах данных сложной структуры, а также при слежении за динамикой и качеством из-

менения данных в заданных информационных областях информационного менеджмента.

Оценка деятельности через визуализацию образов посредством «доски объявлений» мультиагентной системы позволяет создать инструментальные средства «фокусировки внимания», учитывающие символическое мышление системы ИИ и образное мышление ЛПР.

Список литературы

1. Люггер Джордж Ф. Искусственный интеллект: стратегии и методы решения сложных проблем: пер. с англ. – 4-е изд. – М.: Вильямс, 2003. – 864 с.: ил.
2. Рассел С., Норинг П. Искусственный интеллект: современный подход: пер. с англ. – 2-е изд. – М.: Вильямс, 2006. – 1408 с. : ил.
3. Солсо Р. Когнитивная психология. – СПб.: Питер, 2002. – 592 с.: ил. – (Серия «Мастера психологии»).
4. Божич В.И. Концептуальная модель психологии совместной деятельности естественного и искусственного интеллектов / Н.В. Горбатюк, М.Б. Савченко // Известия ТРТУ. Тематический выпуск. Материалы научно-технической конференции «Медицинские информационные системы. МИС-2004». – Таганрог: Изд-во ТРТУ, 2004 – № 6(41) – С. 295–300.
5. Дормашев Ю.Б., Романов В.Я. Психология внимания: учебник. – 3-е изд., испр. – М.: Московский психолого-социальный институт: Флинта, 2002. – 376 с.
6. Линдсей П., Норман Д. Переработка информации у человека (Введение в психологию). – М.: Мир, 1974. – 550 с.
7. Восприятие. Механизмы и модели: сб. статей: пер. с англ. – М.: Мир, 1974. – 367 с.
8. Блум, Ф., Лейзерсон А., Хофстедтер Л. Мозг, разум и поведение: пер с англ. – М.: Мир, 1988. – 248 с.
9. Хьюбел, Д. Глаз, мозг, зрение: пер с англ. – М. : Мир, 1990. – 239 с.
10. Пospelов Д.А. Пушкин В.Н. Мышление и автоматы. – М.: Сов. радио, 1972. – 232 с.
11. Тарасов, В. Б. Моделирование психических образов: как совместить дискретное и непрерывное? // Новости искусственного интеллекта. – 1998. – № 3. – С. 86–100.
12. Тарасов В.Б. Системно-организационный подход в искусственном интеллекте // Программные продукты и системы. – 1997. – № 3. – С. 6–13.

13. Кобринский Б.А. К вопросу о формальном отражении образного мышления и интуиции специалиста слабоструктурированной предметной области // Новости искусственного интеллекта. – 1998. – № 3. – С. 64–76.
14. Фоминых И.Б. Интеграция логических и образных методов отражения информации в системах искусственного интеллекта // Новости искусственного интеллекта. – 1998. – № 3 – С. 76–85.
15. Тарасов В.Б. Агенты, многоагентные системы, виртуальные сообщества: стратегическое направление в информатике и искусственном интеллекте // Новости искусственного интеллекта. – 1998. – № 2. – С. 5–63.
16. Тарасов В.Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. – 352 с. – (Науки об искусственном).
17. Трахтенгерц Э.А. Компьютерная поддержка принятия решений: науч.-практич. изд. Серия «Информатизация России на пороге XXI века». – М.: СИНТЕГ, 1998. – 376 с.

РАЗДЕЛ 8

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

УДК 378.1:004.89

Е.А. Ершова

МОДЕЛЬ ИЕРАРХИЧЕСКОЙ НЕЧЁТКОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЯ В ВУЗЕ

Сложность принятия решения в вузе обусловливается многими факторами, в том числе противоречивостью природы образовательных услуг, которые являются социально-значимыми благами. Поэтому при оценке деятельности вуза необходимо учитывать не только факторы экономического развития, обеспечения максимальной прибыли, но и необходимость выполнения социальных обязательств [1].

При принятии решений в процессе управления сложными системами для обеспечения качества управляющих воздействий, которые бы с максимальной вероятностью обеспечили развитие организации, применяются системы поддержки принятия решений на основе интеллектуальных технологий (рис. 1). Задача заключается в разработке модели системы нечёткого вывода, в которой выходным параметром является решение об изменении плана набора абитуриентов по определённой программе подготовки. Решение заключается в ответе на вопрос, оставить ли контрольные цифры приёма на прежнем уровне, сократить или увеличить.

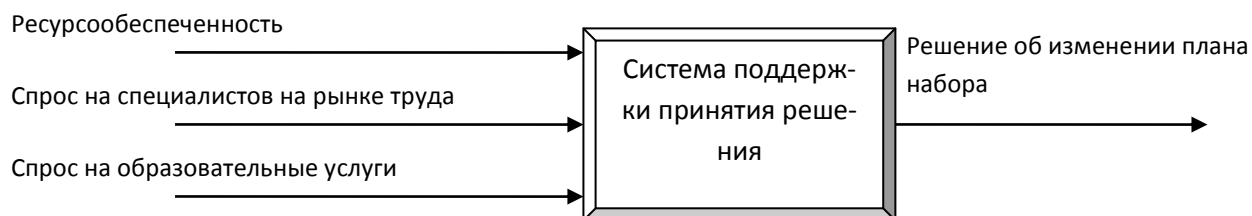


Рис. 1. Модель системы поддержки принятия решения

В случае принятия решения об изменении плана набора абитуриентов учитываются следующие ограничения: институциональные, ресурсные, спрос на образовательные услуги, спрос на выпускников. Эти ограничения являются входными данными при принятии решения, исходящей будет информация о принятом решении.

При принятии решения о планировании набора абитуриентов по специальности необходимо выполнять следующие ограничения:

- обеспечивать набор студентов в разрезе специальностей, направлений и форм обучения в рамках спроса на образовательные услуги;
- обеспечивать выпуск специалистов с учётом спроса на рынке труда;
- в процессе подготовки выполнять требования по обеспечению учебного процесса всеми необходимыми ресурсами должного качества; к этим ресурсам относятся материальные, кадровые, научные и т.п.;
- осуществлять подготовку студентов только при наличии всех необходимых разрешительных и аттестационных документов.

В сформулированной задаче параметры, которые необходимо учитывать при принятии решения, не могут быть сформулированы в числовых, точно измеренных значениях. Например, спрос на рынке образовательных услуг и рынке труда учитывается не в точном измерении, а в некотором диапазоне значений. С другой стороны, достаточно сложно измерить ресурсообеспеченность в рамках подготовки на одном направлении. Это касается и таких параметров, как информационное и материальное обеспечение из расчёта на планируемый контингент. Эти ресурсы можно было бы оценить с точки зрения качества и достаточности по экспертным оценкам, например «кадровое обеспечение на высоком уровне» [3].

Принятие решения в таком случае зависит от ряда нечётких факторов. Перечень этих факторов может быть выстроен в иерархическую зависимость в части ресурсообеспеченности. Ресурсное обеспечение включает в себя следующие виды: кадровое, информационно-методическое, материально-техническое, социальное, обеспечение базами практик (рис. 2).

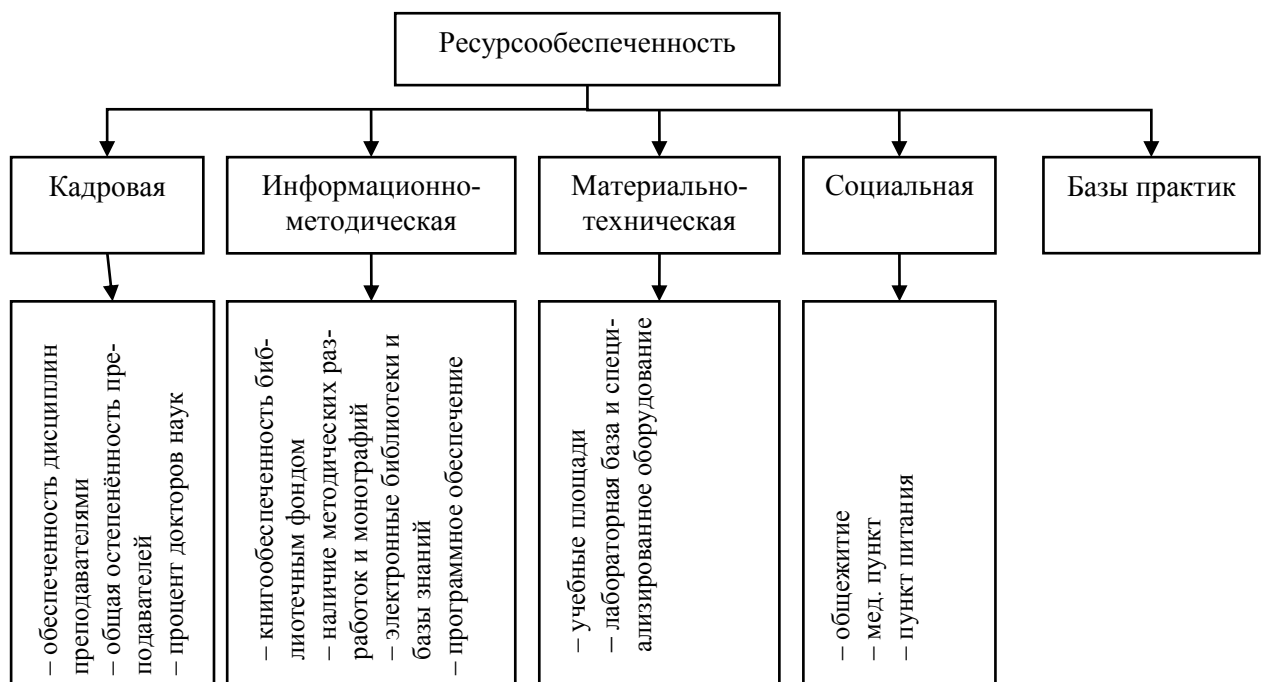


Рис. 2. Классификация параметров ресурсообеспеченности

Данный перечень составлен, исходя из аккредитационных и лицензионных требований для высших учебных заведений. Каждый из перечисленных параметров является комплексным, за исключением обеспеченности базами практик, который может быть оценён только по одному показателю – количеству договоров с предприятиями из расчёта на одного студента.

Таким образом, принятие решения об открытии новой специальности зависит от значения 15 нечётких факторов x_i . Для факторов ресурсообеспеченности определяются лингвистические термы: «выше нормы», «соответствует норме», «ниже нормы»; для факторов спроса – «ниже предложения», «соответствуют предложению», «выше предложения». И для этих термов определяются функции принадлежности.

Само решение об открытии направления также является нечёткой величиной y , для которого определены три лингвистических термина: «сократить», «сохранить», «увеличить».

Поскольку набор абитуриентов, а также и обеспеченность ресурсами часто рассчитываются исходя из количества студенческих групп, термины «увеличить» и «сократить» интерпретируются как изменение плана набора в пределах одной студенческой группы.

Нечёткая иерархическая система в данном случае моделирует зависимость:

$$y = f(x_1, x_2, \dots, x_{15}),$$

где y – исходящая переменная с решением об изменении плана набора; $x_1, x_2, \dots, x_{15}$ – нечёткие параметры, описывающие спрос и ресурсообеспеченность в рамках модели (рис. 1).

Нечёткая система является многофакторной, поэтому описать зависимости в виде совокупности предикатных правил, как требует механизм нечётких выводов, достаточно сложно. Количество правил в системе нечёткого вывода при большом количестве входящих переменных очень велико. Для преодоления «проклятия размерности» необходимо построить иерархию систем нечёткого вывода.

Входами для этой системы будут следующие 15 параметров: x_1 – обеспеченность дисциплин преподавателями; x_2 – общая остепенённость преподавателей; x_3 – процент преподавателей со степенью доктора наук; x_4 – книгообеспеченность библиотечным фондом; x_5 – наличие методических разработок и монографий; x_6 – электронные библиотеки и базы знаний; x_7 – программное обеспечение; x_8 – учебные площади; x_9 – лабора-

торная база и специализированное оборудование; x_{10} – общежитие; x_{11} – мед. пункт; x_{12} – пункт питания; x_{13} – обеспеченность базами практик; x_{14} – спрос на образовательные услуги; x_{15} – спрос на специалистов на рынке труда.

Каждый из указанных параметров является нечётким множеством с функцией принадлежности $\mu_x(t)$, где t – множество допустимых значений критерия.

Построение функций принадлежности может осуществляться прямым или косвенным методами [4]. Прямые методы характеризуются непосредственным описанием функций принадлежности экспертами. Эти методы применимы в случае измеримых значений. Для поставленной задачи применимы косвенные методы, в частности метод статистической обработки экспертной информации [5]. Функция принадлежности строится на основе интервальных оценок группы экспертов. Согласно данной методике на множестве элементов t , на котором определяют функцию принадлежности $\mu_x(t)$, выделяют значения t_j . Каждый из K экспертов заполняет опросник, в котором указывает мнение о наличии у элемента t_j свойств нечёткого множества l_j .

Данная оценка является бинарной и обозначается $b_{i,j}^k \in \{0,1\}$. По результату опроса экспертов степень принадлежности нечёткому множеству l_j рассчитывается следующим образом:

$$\mu_{l_j}(t_j) = \frac{1}{K} \sum_{k=1, K} b_{i,j}^k. \quad (1)$$

По найденным точкам строится кривая, которая описывает лингвистический терм в виде:

$$l_j = \frac{\mu_{l_j}(t_1)}{t_1}, \frac{\mu_{l_j}(t_2)}{t_2}, \dots, \frac{\mu_{l_j}(t_n)}{t_n}. \quad (2)$$

Полученная кривая может быть аппроксимирована какой-либо из известных функций.

Данная иерархическая нечёткая система (рис. 3) будет включать следующие подсистемы:

- $y_1 = f_1(x_1, x_2, x_3)$; кадровое обеспечение;
- $y_2 = f_2(x_4, x_5, x_6, x_7)$; информационно-методическое;
- $y_3 = f_3(x_8, x_9)$; материально-техническое;
- $y_4 = f_4(x_{10}, x_{11}, x_{12}, x_{13})$; социальное;
- $y_5 = f_5(y_1, y_2, y_3, y_4, x_{13})$; общая ресурсообеспеченность.

Итоговая система нечёткого вывода будет в качестве входов получать значения из промежуточных баз знаний y_1, y_2, y_3, y_4 и входящих переменных x_{13}, x_{14}, x_{15} .

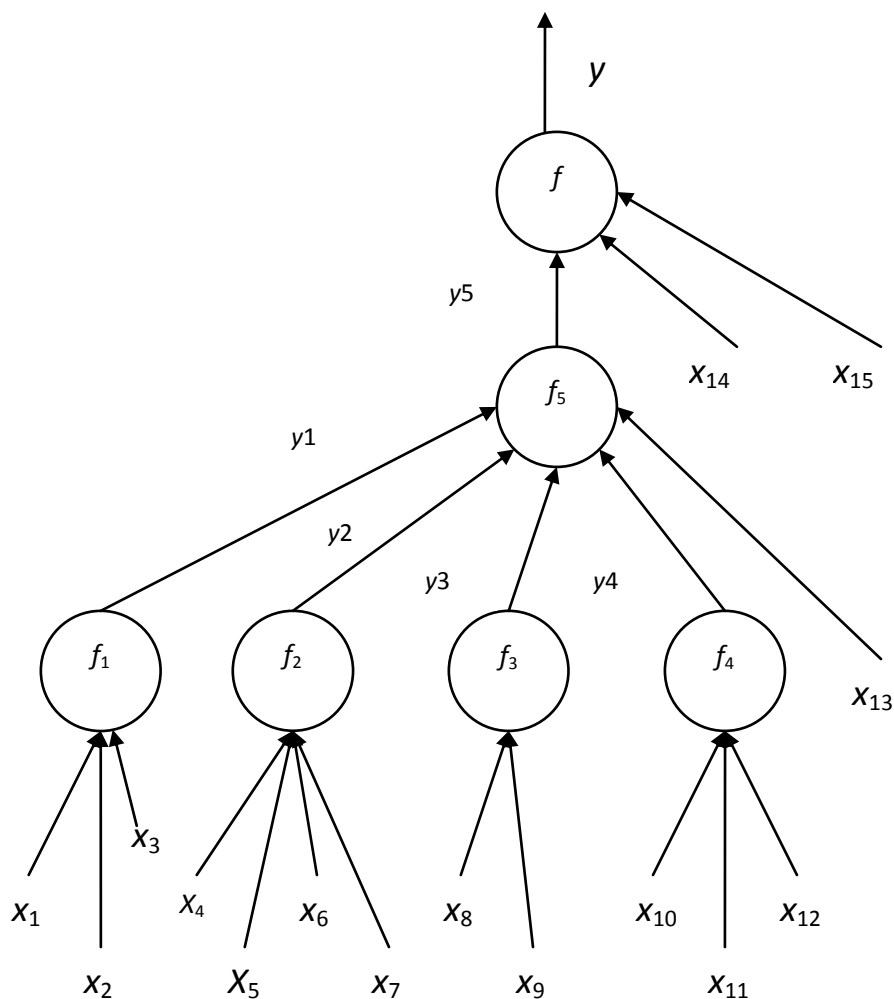


Рис. 3. Иерархическая система нечёткого вывода

Результатами выполнения блоков f_1, f_2, f_3, f_4 являются нечёткие множества, полученные путём пересечения нечётких множеств, описывающих входящие критерии [6]. Так,

$$\begin{aligned}
 y_1 &= x_1 \cap x_2 \cap x_3; \\
 y_2 &= x_4 \cap x_5 \cap x_6 \cap x_7; \\
 y_3 &= x_8 \cap x_9; \\
 y_4 &= x_{10} \cap x_{11} \cap x_{12}; \\
 y_5 &= y_1 \cap y_2 \cap y_3 \cap y_4 \cap x_{13}.
 \end{aligned}
 \tag{3}$$

При этом необходимо отметить различную степень влияния каждого фактора на результат решения общей задачи. Этот факт учитывается путём добавления весов α_i для каждого входящего аргумента, тогда результативное нечёткое множество y определяется следующим образом:

$$\begin{aligned}\mu_{y_1}(t) &= \alpha_1 \mu_{x_1}(t) \wedge \alpha_2 \mu_{x_2}(t) \wedge \alpha_3 \mu_{x_3}(t) \\ \mu_{y_2}(t) &= \alpha_4 \mu_{x_4}(t) \wedge \alpha_5 \mu_{x_5}(t) \wedge \alpha_6 \mu_{x_6}(t) \wedge \alpha_7 \mu_{x_7}(t) \\ \mu_{y_3}(t) &= \alpha_8 \mu_{x_8}(t) \wedge \alpha_9 \mu_{x_9}(t) \\ \mu_{y_4}(t) &= \alpha_{10} \mu_{x_{10}}(t) \wedge \alpha_{11} \mu_{x_{11}}(t) \wedge \alpha_{12} \mu_{x_{12}}(t) \\ \mu_{y_5}(t) &= \alpha_{14} \mu_{y_1}(t) \wedge \alpha_{15} \mu_{y_2}(t) \wedge \alpha_{16} \mu_{y_3}(t) \wedge \alpha_{17} \mu_{y_4}(t) \wedge \alpha_{13} \mu_{x_{13}}(t)\end{aligned}, \quad (4)$$

где α_i – коэффициенты важности критерия.

Коэффициент α_i определяется путём попарного их сравнения и построения итогового вектора коэффициентов по методу Саати [7].

В предложенной модели иерархической системы блок $f(y_5, x_{14}, x_{15})$ реализован в виде системы нечёткого вывода по алгоритму Мамдани.

Для работы любой системы нечёткого вывода характерны следующие этапы: фазификация, нечёткий вывод, композиция, дефазификация [8].

На первом этапе – фазификации – функции принадлежности применяются к их фактическим значениям. Фактическое значение переменной y_5 будет получено как результат пересечения нечётких множеств, описывающих входящие критерии. Результат будет получен в виде степеней принадлежности значения лингвистическим термам $A_1(y_5)$, $A_2(y_5)$, $A_3(y_5)$. Значения для x_{14} , x_{15} будут задаваться экспертом. Для каждого фактического значения переменных будет определена степень истинности исходя из функций принадлежности для трёх термов («ниже предложения», «соответствуют предложению», «выше предложения»): $B_1(x_{14})$, $B_2(x_{14})$, $B_3(x_{14})$, $C_1(x_{15})$, $C_2(x_{15})$, $C_3(x_{15})$.

Согласно алгоритму Мамдани, на втором этапе нечёткого вывода происходит «отсечение» для предпосылок каждого из правил с использованием операции логического минимума. Механизм нечётких выводов в своей основе имеет базу правил, связывающих значения переменных y_5 , x_{14} , x_{15} с выходной переменной y , которая представлена в таблице.

На третьем этапе – композиции – объединяются усечённые на втором этапе функции принадлежности с помощью операции логического максимума. В результате получается комбинированная функция принадлежности $\mu_\Sigma(y)$, соответствующая логическому выводу для выходной переменной y .

База правил для системы нечёткого вывода

Ресурсообеспеченность (y_5)	Спрос на образовательные услуги (x_{14})	Спрос на специалистов на рынке труда (x_{15})	Изменение плана набора (y)
выше нормы	превышает предложение	превышает предложение	увеличить
соответствует норме	превышает предложение	превышает предложение	увеличить
ниже нормы	превышает предложение	превышает предложение	сохранить
выше нормы	соответствует предложению	превышает предложение	увеличить
соответствует норме	соответствует предложению	превышает предложение	сохранить
ниже нормы	соответствует предложению	превышает предложение	сохранить
выше нормы	ниже предложения	превышает предложение	сохранить
соответствует норме	ниже предложения	превышает предложение	сохранить
ниже нормы	ниже предложения	превышает предложение	сократить
выше нормы	превышает предложение	соответствует предложению	увеличить
соответствует норме	превышает предложение	соответствует предложению	увеличить
ниже нормы	превышает предложение	соответствует предложению	сократить
выше нормы	соответствует предложению	соответствует предложению	сохранить
соответствует норме	соответствует предложению	соответствует предложению	сохранить
ниже нормы	соответствует предложению	соответствует предложению	сохранить
выше нормы	ниже предложения	соответствует предложению	сохранить
соответствует норме	ниже предложения	соответствует предложению	сократить
ниже нормы	ниже предложения	соответствует предложению	сократить
выше нормы	превышает предложение	ниже предложения	увеличить
соответствует норме	превышает предложение	ниже предложения	увеличить
ниже нормы	превышает предложение	ниже предложения	сократить
выше нормы	соответствует предложению	ниже предложения	сохранить
соответствует норме	соответствует предложению	ниже предложения	сохранить
выше нормы	ниже предложения	ниже предложения	сократить
ниже нормы	ниже предложения	ниже предложения	сократить
соответствует норме	ниже предложения	ниже предложения	сократить
ниже нормы	ниже предложения	ниже предложения	сократить

На четвёртом этапе – дефазификации – находится чёткое значение выходной переменной. Один из методов его нахождения – центроидный, в котором поиск чёткого значения y_0 определяется по формуле:

$$y_0 = \frac{\int y\mu_{\Sigma}(y)dy}{\int \mu_{\Sigma}(y)dy}. \quad (5)$$

Таким образом, нами представлена модель системы поддержки принятия решения на основе нечётких множеств и нечёткого вывода по алгоритму Мамдани. Применение данной системы позволит повысить качество и обоснованность принимаемых решений при планировании набора абитуриентов в вузе в разрезе специальностей и направлений с учётом трёх основных параметров – ограничений: наличия ресурсов на подготовку, спроса на образовательные услуги и прогнозируемого спроса на рынке труда.

Список литературы

1. Береза А.Н., Ершова Е.А. Применение методов поддержки принятия решений в задачах реструктуризации вуза // Открытое образование. – 2010. – № 4. – С. 91–101.
2. Новиков Д.А. Модели и механизмы управления развитием региональных образовательных систем (концептуальные положения). – М.: ИПУ РАН, 2001. – 83 с.
3. Ершова Е.А. Процедура принятия решения в управлении вузом на основе теории нечётких множеств // Управление знаниями и технологии – С. 179–182.
4. Нечёткие множества в моделях управления и искусственного интеллекта / под ред. Д.А. Поспелова. – М.: Гл. ред. физ.-мат. лит., 1986. – 312 с.
5. Штовба С.Д. Проектирование нечётких систем средствами MATLAB. – М.: Горячая линия – Телеком, 2007. – 288 с.
6. Беллман Р., Заде Л. Принятие решений в расплывчатых условиях // Вопросы анализа и процедуры принятия решений. – М.: Мир, 1976. – С. 172–215.
7. Саати Т.Л. Принятие решений при зависимостях и обратных связях. – М.: Изд-во ЛКИ, 2007. – 357 с.
8. Круглов В.В., Дли М.И., Голунов Р.Ю. Нечёткая логики и искусственные нейронные сети: учеб. пособие. – М.: Изд-во Физико-математической литературы, 2001. – 224 с.

СВЕДЕНИЯ ОБ АВТОРАХ

Алабут А.В.	Ростовский государственный медицинский университет, г. Ростов-на-Дону
Алексеев А.М.	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, г. Санкт-Петербург
Алепко А.В.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Андрющенко И.В.	Ростовский технологический институт сервиса и туризма (РТИСТ ЮРГУЭС), г. Ростов-на-Дону
Асцатуров Ю.Г.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Бахман Е.А.	Волгоградский государственный технический университет, г. Волгоград
Бегляров В.В.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Безуглов Д.А.	Ростовский технологический институт сервиса и туризма (РТИСТ ЮРГУЭС), г. Ростов-на-Дону
Бекяшева З.С.	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, г. Санкт-Петербург
Божич В.И.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Валюкевич Ю.А.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Васильев В.Н.	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, г. Санкт-Петербург
Востриков А.А.	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, г. Санкт-Петербург
Галушкин Н.Е.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Галушкина Н.Н.	Ростовский технологический институт сервиса и туризма (РТИСТ ЮРГУЭС), г. Ростов-на-Дону
Гладков Л.А.	Таганрогский технологический институт Южного федерального университета (ТТИ ЮФУ), г. Таганрог

Дубовсков В.В.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Ершова Е.А.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Зеленский А.А.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Ионин В.К.	Российский государственный социальный университет, г. Москва
Кизим А.В.	Волгоградский государственный технический университет, г. Волгоград
Клименко Н.Б.	Ростовский институт (филиал) Российского государственного торгово-экономического университета, г. Ростов-на-Дону
Козаченко П.Н.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Козоброд А.В.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Коровин Е.Н.	Кемеровский государственный университет, г. Кемерово
Коротков А.В.	Международный центр теоретической физики, г. Новочеркасск
Куркина М.В.	Югорский государственный университет, г. Ханты-Мансийск
Лисовцова А.Е.	Таганрогский технологический институт Южного федерального университета (ТТИ ЮФУ), г. Таганрог
Львова М.А.	Алтайская государственная педагогическая академия, Барнаул
Медведев Д.В.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Мешков В.Е.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Мешкова Е.В.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Наумов И.И.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Никонорова Ю.В.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск

Орлов В.В.	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, г. Санкт-Петербург
Павлов А.В.	Санкт-Петербургский государственный университет информационных технологий, механики и оптики, г. Санкт-Петербург
Пархоменко Н.Г.	Ростовский технологический институт сервиса и туризма (РТИСТ ЮРГУЭС), г. Ростов-на-Дону
Подстрелов А.Г.	ЗАО НТ-МДТ, г. Зеленоград
Попова З.И.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Савченко М.Б.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Семенов В.В.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Стороженко А.С.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Токарев А.А.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Толстунов О.Г.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Трясоруков А.И.	Ростовский институт (филиал) Российского государственного торгово-экономического университета, г. Ростов-на-Дону
Ханжонков Ю.Б.	Южно-Российский государственный университет экономики и сервиса, г. Шахты
Чураков В.С.	Волгодонский институт сервиса (филиал) Южно-Российского государственного университета экономики и сервиса (ВИС ЮРГУЭС), г. Волгодонск
Швидченко С.А.	Ростовский технологический институт сервиса и туризма (РТИСТ ЮРГУЭС), г. Ростов-на-Дону

.

Научное издание

**ИНФОРМАЦИОННЫЕ СИСТЕМЫ
И ТЕХНОЛОГИИ.
ТЕОРИЯ И ПРАКТИКА**

Сборник научных трудов

Ответственный за выпуск Н.В. Ковбасюк

ИД № 06457 от 19.12.01 г. Издательство ЮРГУЭС.

Подписано в печать 4.07.11 г.

Формат бумаги 60х90/16. Усл. печ. л. 17,75.

Тираж 65 экз. Заказ № 380.

ПЛД № 65-175 от 05.11.99 г.

Типография Издательства ЮРГУЭС.

346500, г. Шахты, Ростовская обл., ул. Шевченко, 147

