

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
БУРЯТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

В Е С Т Н И К  
БУРЯТСКОГО ГОСУДАРСТВЕННОГО  
УНИВЕРСИТЕТА

Выпуск 9

МАТЕМАТИКА, ИНФОРМАТИКА

*Журнал включен Высшей аттестационной комиссией в Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук*



ИЗДАТЕЛЬСТВО

Улан-Удэ

2012

*Редакционный совет «Вестника БГУ»*

*С.В. Калмыков*, чл.-кор. РАО, д-р пед. наук, проф. (председатель);  
*А.С. Булдаев*, д-р физ.-мат. наук, проф. (зам. председателя); *Н.Н. Татарникова*  
(зам. председателя, директора Издательства); *М.В. Бадмаева*, д-р филос. наук,  
доц.; *Т.С. Базарова*, д-р пед. наук, доц.; *Д.И. Бураев*, д-р ист. наук, проф.;  
*А.В. Гаськов*, д-р пед. наук, проф.; *Н.Ж. Дагбаева*, д-р пед. наук, проф.;  
*Ц.З. Доржиев*, д-р биол. наук, проф.; *С.С. Имixelова*, д-р филол. наук, проф.;  
*Л.П. Ковалева*, канд. филол. наук, проф.; *К.Б-М. Митупов*, д-р ист. наук,  
проф.; *И.И. Осинский*, д-р филос. наук, проф.; *М.Н. Очиров*, д-р пед. наук,  
проф.; *В.В. Хахинов*, д-р хим. наук, проф. *В.Е. Хитрихеев*, д-р мед. наук, проф.

*Редакционная коллегия выпуска*

*А.С. Булдаев*, д-р физ.-мат. наук, проф. (гл. редактор); *С.Н. Васильев*, акад. РАН;  
*В.И. Гурман*, д-р техн. наук, проф.; *А.Д. Миждон*, д-р техн. наук, проф.;  
*В.Б. Цыренова*, д-р пед. наук, доц.; *Д.Ш. Ширапов*, д-р физ.-мат. наук, проф.;  
*В.И. Антонов*, канд. физ.-мат. наук, доц.; *Т.Г. Дармаев*, канд. физ.-мат. наук, доц.;  
*И.К. Шаранхаев*, канд. физ.-мат. наук, доц.; *Г.А. Шишкин*, канд. физ.-мат. наук,  
проф.; *И.Б. Юмов*, канд. физ.-мат. наук, доц.; *В.В. Кибирев*, канд. физ.-мат. наук, проф.

В Е С Т Н И К  
БУРЯТСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Выпуск 9

МАТЕМАТИКА, ИНФОРМАТИКА



# 1. Управляемые системы и методы оптимизации

УДК 519.71

© О.И. Дранко, Г.Н. Яковенко

## ОПТИМИЗАЦИЯ ДИВИДЕНДНОЙ ПОЛИТИКИ<sup>1</sup>

Статья посвящена поиску оптимальной дивидендной политики для максимизации стоимости бизнеса. Методом оптимизации выступают теоретико-групповой анализ и принцип максимума Понтрягина. Оптимальное решение предполагает этапы накопления и потребления. Полученный результат сравнивается с практикой российских и американских компаний.

**Ключевые слова:** Стоимость бизнеса, оптимальное управление, дивидендная политика, групповой анализ.

*O.I. Dranko, G.N. Yakovenko*

## THE OPTIMIZATION OF DIVIDEND POLICY

The article is devoted to finding the optimal dividend policy to maximize business value. The method of optimization of theoretic analysis of a Group and the maximum principle of Pontryagin. The optimal solution involves the stages of accumulation and consumption. This result is compared with the practice of Russian and American companies.

**Keywords:** value of a company, optimal control, dividend policy, group analysis.

### Введение

На новом этапе развития России существенно меняется целевой показатель: предстоит задача формирования и увеличения стоимости бизнеса (для Советского периода формальным показателем являлась прибыль, для бурных 1990-х – денежный поток). Эта задача существенно более сложная, так как требует управления гораздо более широким перечнем параметров.

В данной работе рассматривается модель стоимости по дисконтированным дивидендам в прогнозный период и поиск оптимального решения по выплате дивидендов.

### 1. Постановка задачи. Математическая модель

Рассмотрим упрощенную модель развития бизнеса без заемных средств за счет реинвестирования части прибыли после налогов, описываемую уравнением (конечно-разностный вариант)

$$\Delta A = u\pi = u(mA - F) \quad (1.1)$$

где  $A$  — активы,  $\pi$  — чистая прибыль,  $u$  — доля чистой прибыли на накопление (играет роль управления),  $m$  — маржинальная рентабельность активов,  $F$  — постоянные затраты.

Стоимость бизнеса для акционеров определяется дисконтированным денежным потоком для акционеров [1], т.е. выводом чистой прибыли для потребления (дивиденды). Предполагается, что менеджмент должен максимизировать стоимость для акционеров. Рассмотрим модель стоимости на прогнозном периоде (непрерывный вариант модели (1.1)):

$$V = \int_0^T \frac{(1-u)(mA - F)}{(1+r)^t} dt = \int_0^T f_0 dt \rightarrow \max \quad (1.2)$$

$$\frac{dA}{dt} = u(mA - F), \quad 0 \leq u \leq 1, \quad (1.3)$$

где  $f_0$  — функция потребления (дивиденды),  $r$  — ставка дисконтирования. Далее предполагается  $r = 0$ ,  $m > 0$ ,  $F > 0$ . Для компактности введём следующие обозначения  $x_1 = A$ ,  $x_2 = -V$  (см. (1.2), (1.3)). В этих обозначениях оптимальная задача примет вид ( $\dot{x} = dx/dt$ )

$$\begin{aligned} \dot{x}_1 &= (mx_1 - F)u, \\ \dot{x}_2 &= -(mx_1 - F)(1-u), \end{aligned} \quad (1.4)$$

<sup>1</sup> Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 10-01-00228).

$$0 = t_0 \leq t \leq T, \quad mx_1^0 > F > 0, \quad x_2^0 = 0, \quad (1.5)$$

$$0 \leq u \leq 1, \quad (1.6)$$

$$x_2(T) \rightarrow \min. \quad (1.7)$$

## 2. Теоретико-групповой анализ модели

Подвергнем систему (1.4) теоретико-групповому анализу [2]. Уравнениям (1.4) соответствует семейство операторов дифференцирования по  $t$  в силу этой системы:

$$X(u) = uX_1 + X_2,$$

где обозначено

$$X_1 = (mx_1 - F) \left( \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2} \right), \quad X_2 = -(mx_1 - F) \frac{\partial}{\partial x_2}. \quad (2.1)$$

Вычисление коммутатора приводит к результату

$$[X_1, X_2] = -m(mx_1 - F) \frac{\partial}{\partial x_2} = mX_2.$$

Таким образом, операторы (2.1) есть базис алгебры Ли с структурными постоянными<sup>2</sup>  $C_{12}^2 = m$ , а система (1.4) является групповой. По уравнениям системы вычисляется 2-параметрическая группа ( $v_1, v_2$  — параметры):

$$\begin{aligned} x_1 &= x_1^0 - \frac{F}{m} e^{v_1} - 1, \\ x_2 &= x_2^0 - mx_1^0 - F v_2 \frac{e^{v_1} - 1}{v_1} \end{aligned} \quad (2.2)$$

— группа сдвигов вдоль решений системы (1.4). Выбор управления  $u(t)$  и интервала  $[0, T]$  определяет преобразование пространства состояний: начальные состояния  $x_i^0$  решения  $x_i(t)$  переносят в состояния  $x_i(T)$ . Описанное преобразование есть преобразование группы (2.2) при некотором выборе параметров  $v_1, v_2$ .

Так как количество (два) параметров в группе (2.2) совпадает с размерностью (два) пространства состояний, то система (1.4) является не только групповой, но в терминологии [2] и  $L$ -системой.  $L$ -системы размерности  $n$  кроме  $n$ -параметрической группы сдвигов вдоль решений (в данном случае (2.2)) допускает  $n$ -параметрическую группу симметрий. Для её подсчёта нужно сначала вычислить коэффициенты соответствующего группе оператора

$$Y = \eta_1(x_1, x_2) \frac{\partial}{\partial x_1} + \eta_2(x_1, x_2) \frac{\partial}{\partial x_2}. \quad (2.3)$$

Оператор симметрий (2.3) связан с операторами (2.1) следующим образом  $X_i, Y = 0$ . Раскрыв коммутаторы, получим следующие уравнения для функций  $\eta_1(x_1, x_2)$  и  $\eta_2(x_1, x_2)$ :

$$\begin{aligned} (mx_1 - F) \left( \frac{\partial \eta_1}{\partial x_1} + \frac{\partial \eta_1}{\partial x_2} \right) - m\eta_1 &= 0, \\ (mx_1 - F) \left( \frac{\partial \eta_2}{\partial x_1} + \frac{\partial \eta_2}{\partial x_2} \right) - m\eta_2 &= 0, \\ -(mx_1 - F) \frac{\partial \eta_1}{\partial x_2} &= 0, \\ -(mx_1 - F) \frac{\partial \eta_2}{\partial x_2} + m\eta_1 &= 0. \end{aligned}$$

<sup>2</sup> Здесь и далее приводятся только ненулевые постоянные  $C_{ij}^k$  при  $i < j$ .

Уравнения без труда решаются, что приводит к общему решению ( $c_1, c_2$  — произвольные постоянные):

$$\eta_1 = -c_1(mx_1 - F), \quad \eta_2 = -c_1mx_2 + c_2. \quad (2.4)$$

Полагая  $c_1 = 1, c_2 = 0$  и  $c_1 = 0, c_2 = 1$ , приходим к двум операторам симметрий

$$Y_1 = -(mx_1 - F)\frac{\partial}{\partial x_1} - mx_2\frac{\partial}{\partial x_2}, \quad Y_2 = \frac{\partial}{\partial x_2} \quad (2.5)$$

— базису алгебры Ли операторов симметрий. Отметим, что для приведённых операторов выполняется  $[Y_1, Y_2] = mY_2$ , то есть алгебра Ли операторов симметрий имеет те же структурные постоянные  $C_{12}^2 = m$ , что и алгебра сдвигов вдоль решений.

Для перехода от операторов (2.5) к уравнениям группы симметрий нужно по коэффициентам (2.4) оператора построить систему

$$\begin{aligned} \frac{d\hat{x}_1}{d\tau} &= -c_1(m\hat{x}_1 - F), & \hat{x}_1(0) &= x_1, \\ \frac{d\hat{x}_2}{d\tau} &= -c_1m\hat{x}_2 + c_2, & \hat{x}_2(0) &= x_2, \end{aligned}$$

решить её и ввести обозначения  $c_i\tau = \tau_i$ :

$$\begin{aligned} \hat{x}_1 &= x_1 + F\frac{1 - e^{-m\tau_1}}{m}, \\ \hat{x}_2 &= x_2e^{-m\tau_1} + \tau_2\frac{1 - e^{-m\tau_1}}{m\tau_1}. \end{aligned} \quad (2.6)$$

Польза от группы симметрий заключается, в частности, в том, что если любое решение  $x_i(t)$  системы (1.4) при некотором допустимом управлении  $u(t)$  подставить в уравнения группы (2.6), то получится 2-параметрическое семейство решений  $\hat{x}_i(t, \tau_1, \tau_2)$ , соответствующих тому же управлению  $u(t)$ .

### 3. Решение оптимальной задачи

При решении оптимальной задачи (1.4)–(1.7) будем следовать формализму принципа максимума Л.С. Понтрягина [3]. Задаче соответствует функция Гамильтона

$$H = p_1(mx_1 - F) + p_2(mx_1 - F)u - p_2(mx_1 - F), \quad (3.1)$$

уравнения Гамильтона

$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial x_1} = -m(p_1 + p_2)u - p_2, \\ \dot{p}_2 &= -\frac{\partial H}{\partial x_2} = 0, \end{aligned} \quad (3.2)$$

условия трансверсальности (1.7):

$$p_1(T) = -\frac{\partial x_2}{\partial x_1}\Big|_{t=T} = 0, \quad p_2(T) = -\frac{\partial x_2}{\partial x_2}\Big|_{t=T} = -1. \quad (3.3)$$

Вследствие того, что система (1.4) является  $L$ -системой, для системы (1.4), (3.2) по коэффициентам (2.4) оператора симметрий (2.5) строится семейство первых интегралов [2]

$$w(c_1, c_2) = p_1(-c_1mx_1 + F) + p_2(-c_1mx_2 + c_2).$$

В семействе выделяется базис

$$\begin{aligned} w_1 &= w(-1, 0) = p_1(mx_1 - F) + p_2mx_2, \\ w_2 &= w(0, 1) = p_2. \end{aligned} \quad (3.4)$$

Полагая в (3.4)  $t = T$ , с учётом (3.3) получим выражения (введено обозначение  $x_2(T) = \bar{x}_2$ )

$$\begin{aligned} p_1(mx_1 - F) + p_2mx_2 &= -m\tilde{x}_2, \\ p_2 &= -1, \end{aligned}$$

из которых следует

$$p_1(mx_1 - F) = m(x_2 - \tilde{x}_2), \quad p_2 = -1.$$

Учёт этих зависимостей в (3.1) приводит к формуле

$$H = m(x_2 - \tilde{x}_2) - mx_1 + F u + mx_1 - F$$

или

$$H = m(x_2 - x_1 + a)u + mx_1 - F, \tag{3.5}$$

обозначено  $a = \frac{F}{m} - \tilde{x}_2$ . В силу принципа максимума с учётом ограничения (1.6) для оптимального управления получаем необходимое условие

$$u = \begin{cases} 0 & \text{при } x_2 - x_1 + a < 0, \\ 1 & \text{при } x_2 - x_1 + a > 0. \end{cases} \tag{3.6}$$

Из условия видно, что по одну сторону прямой  $x_2 - x_1 + a = 0$  управление принимает значение  $u = 0$ , а по другую  $u = 1$ . Рассмотрим варианты в зависимости от начальных условий.

1.  $x_1^0 > a$ , В силу (3.6) при  $t=0$  выполняется  $u=0$ , а в силу уравнений (1.4) при  $t>0$ :  $x_1(t) = x_1^0$ ,  $x_2(t) = -(mx_1^0 - F)t$ , то есть условие, приводящее к  $u=0$  продолжает выполняться, и к моменту  $t=T$  функционал достигает своего минимального значения  $x_2 = -(mx_1^0 - F)T$ .

2.  $x_1^0 < a$ ,  $x_2^0 = 0$ . В силу (3.6) при  $t=0$  выполняется  $u=1$ , а в силу уравнений (1.4) при  $t>0$ :  $mx_1 - F = (mx_1^0 - F)e^{mt}$ ,  $x_2(t) = 0$ . Пусть в момент  $(0 \leq t_1 \leq T)$  произошло переключение. При  $t \geq t_1$  справедливо  $u=0$  и в силу уравнений (1.4):

$$mx_1 - F = (mx_1^0 - F)e^{mt_1}, \quad x_2(t, t_1) = -(mx_1^0 - F)e^{mt_1}(t - t_1).$$

Спрашивается, в какой момент  $t_1$  ( $0 \leq t_1 \leq T$ ) должно произойти переключение, чтобы функционал

$$x_2(T, t_1) = -(mx_1^0 - F)e^{mt_1}(T - t_1) \tag{3.7}$$

достиг своего минимального значения. Исследование на экстремум приводит к соотношению

$$\frac{dx_2(T, t_1)}{dt_1} = (mx_1^0 - F)e^{mt_1} \{-m(T - t_1) + 1\} = 0$$

и к нужному результату

$$t_1 = T - \frac{1}{m}. \tag{3.8}$$

Для того, чтобы результат не покинул границы  $0 \leq t_1 \leq T$  требуется для конечного момента выполнения

$$mT \geq 1. \tag{3.9}$$

В противном случае минимум будет достигаться на границе  $t_1 = 0$ , что соответствует варианту 1. Подставим результат (3.8) в функционал (3.7)

$$x_2(T, t_1) = -(mx_1^0 - F) \frac{1}{m} e^{mT-1}.$$

Несложный анализ показывает, что это значение при условии (3.9) лучше, чем граничное значение  $x_2 = -(mx_1^0 - F)T$ .

#### 4. Содержательное обсуждение результатов

Практика дивидендной политики для 2622 компаний США по итогам 2009 года, котирующихся на бирже [4], приведена в таблице 4.1.

Таблица 4.1

Данные компаний США, 2009 г.

Категория надежности	A++	A+	A	B++	B+	B
Количество компаний	52	73	257	526	902	812
Маржинальная рентабельность активов <sup>3</sup> m, %	41	41	37	35	31	30
Доля чистой прибыли на накопление и, %	63	62	58	54	63	68

Оптимальный режим для развития компании на 10 лет выглядит следующим образом:

- интенсивное развитие в течение  $t_1 = \max(0, T - 1/m)$ . Для данных компаний США это составляет 7 лет;

- последние 3 года вся чистая прибыль направляется на дивиденды.

Длительность периода развития зависит от срока анализа и маржинальной рентабельности активов. Для низкорентабельных компаний этап начального развития сокращен, целесообразен вывод прибыли.

Около 2/3 прибыли компаний США направляется на развитие. Такая практика свидетельствует о компромиссной дивидендной политике (между получением дивидендного денежного потока и реинвестированием), а не оптимизацией расчетной стоимости компании. Данная статистика может свидетельствовать о том, что большинство компаний США находятся в «постпрогнозном» (инерционном) росте.

Выплаты дивидендов в России находятся на относительно невысоком уровне. Основными причинами являются реинвестирование прибыли в развитие (оптимально для быстрого развития) и оптимизация налогов (на прибыль, на дивиденды<sup>4</sup>).

В России выплаты дивидендов определяются скорее практиками корпоративного управления для компаний на фондовой бирже, чем оптимизацией богатства акционеров. Большинство российских компаний платит дивиденды в размере 1-5% чистой прибыли, некоторые (Лукойл, Сбербанк) – 15-20%, подавляющее меньшинство (МТС) – 60-90%. Дивидендная доходность 10-15% (к цене акций) считается редкой и очень хорошей.

#### Заключение

Исходя из представленного в работе исследования видно, что одним из основных параметров управления стоимостью компании является коэффициент реинвестирования прибыли.

Результаты моделирования показывают целесообразность активного развития на начальном этапе и вывода прибыли на этапе зрелости компании.

Оптимизационная дивидендная политика компаний в прогнозный период для максимизации акционерной стоимости не подтверждается практикой (фактическими данными) крупнейших компаний США.

Данный результат показывает или необходимость уточнения модели (что будет проходить из-за уточнения различных факторов), или инерционность (компромиссность, а не оптимальность) действий менеджмента.

Для дальнейших исследований целесообразно развитие модели:

- Анализ влияния ставки дисконтирования. Ставка дисконтирования принята нулевой для упрощения получения оптимального решения. Дисконтирование будущих денежных потоков только усилит эффект ускоренного потребления.

- Анализ заемной политики, влияния структуры капитала на стоимость компании и оптимальную дивидендную политику.

<sup>3</sup> Gross Income / Total Assets.

<sup>4</sup> Ставка налога на дивиденды 9%, кроме организаций, владеющих не менее 50% уставного капитала не менее года.

**Литература**

1. Коупленд Т., Колер Т., Мурин Д. Стоимость компаний: оценка и управление. М.: Олимп-Бизнес, 2005. – 576 с.
2. Яковенко Г.Н. Теория управления регулярными системами. — М.: БИНОМ. Лаборатория знаний, 2008. – 264 с.
3. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М.: Наука, 1961. – 394 с.
4. URL: <http://valueline.com/>

*Олег Иванович Дранко*, канд. физ.-мат. наук, доц. каф. инновационного менеджмента Московского физико-технического института (МФТИ), тел. (495) 576-4022, e-mail: [olegdranko@gmail.com](mailto:olegdranko@gmail.com).

*Геннадий Николаевич Яковенко*, д-р физ.-мат. наук, проф. каф. теоретической механики Московского физико-технического института (МФТИ), тел. (495) 576-5733, e-mail: [yakovenkog@gmail.com](mailto:yakovenkog@gmail.com).

*Oleg Ivanovich Dranko*, candidate of physical and mathematical sciences, associate professor of the Department of innovation management of Moscow institute of physics and technology (MIPT).

*Gennady Nikolaevich Yakovenko*, doctor of physical and mathematical sciences, professor of the Department of theoretical mechanics of Moscow institute of physics and technology (MIPT).

УДК 519.6

© *А.И. Колмакова*

## **МНОГОМЕТОДНАЯ ОПТИМИЗАЦИЯ УПРАВЛЕНИЯ В ЭКОНОМИЧЕСКОЙ МОДЕЛИ ВЫБОРА НАЛОГОВОЙ СТАВКИ**

На основе задачи максимизации непроизводственного потребления в рамках нелинейной однопродуктовой модели с производственной функцией Кобба – Дугласа и однопродуктовой макроэкономической модели производственного предприятия с внесенными поправками, связанными с непосредственным учетом налоговых отчислений, формулируется математическая постановка задачи в виде системы дифференциальных уравнений, описывающих налогообложение прибыли предприятий и рассматривается задача поиска оптимального программного управления.

**Ключевые слова:** Оптимальное управление, динамическое программирование, налог на прибыль, модель деятельности предприятия.

*A.I. Kolmakova*

## **MULTIMETHOD OPTIMIZING OF CONTROL IN THE ECONOMIC MODEL OF THE CHOICE OF TAX RATE**

On the basis of the maximization problem of non-production consumption in the nonlinear single-product model with the production function Cobb – Douglas and the macroeconomic model of a single-production facility with corrections which are related to direct accounting of tax charges, it is formulated a mathematical problem as a system of differential equations that describes the taxing of enterprise profits and is considered the problem of optimal software control seeking.

**Keywords:** Optimal control, dynamic programming, profits tax, enterprise activities model.

### **Введение**

Появление налогов было обусловлено необходимостью удовлетворять растущие потребности финансового хозяйства государства. Процесс взимания налогов осуществляется государством и является его функцией. Способ, характер и масштабы мобилизации денежных ресурсов в бюджете государства зависят от стадии экономического развития общества. При этом минимальный размер формируемых денежных фондов, состоящих в большей степени из налоговых платежей, определяется суммой расходов государства на исполнение его функций: социальной, экономической, оборонной и др.

Система налогообложения является одним из важнейших факторов, влияющих на решение об инвестировании, особенно если речь идет о прямых иностранных инвестициях. В целом на принятие решений об инвестировании оказывают влияние как налоговые, так и неналоговые факторы, при помощи которых осуществляется сравнение потенциальной прибыли от инвестиций и риска. Обычно компании, при выборе того или иного проекта, принимают во внимание различие в законных ставках налога, направляя доходы в страны с низким налогообложением [1].

Оптимальность налогообложения принято оценивать с точки зрения общего эффекта для благосостояния общества и с точки зрения выгод для определенного налогоплательщика (экономического агента). Именно от характера и сущности проводимой государством и его органами общей налогово-бюджетной политики зависит, определять ли налоги как «бремя, «зло» или как способ (эффективный в большей или меньшей степени) перераспределения материальных благ и финансирования общих объективных потребностей общества.

### **1. Многометодные алгоритмы расчета оптимального управления**

Предлагается вычислительная технология расчёта оптимального управления, основанная на применении нескольких процессоров для поиска численного решения задачи несколькими методами оптимизации одновременно. В соответствии с этой технологией после нахождения очередного приближения все методы оцениваются по полученному приращению функционала, и из них выбирается наиболее эффективный метод для продолжения оптимизации, а полученное этим методом приближение передается остальным методам в качестве начального для выполнения следующей итерации.

**1.1. Градиентные методы для решения задач без ограничений на управление**

В методах градиентного типа трудоемкую операцию расчета градиента, требующую интегрирования сопряженной системы, следует выполнить только один раз, а затем использовать найденный градиент в итерационных формулах всех методов. Вычислительные затраты на выполнение одного шага многометодного алгоритма в этом случае значительно сократятся. Кроме того, реализация шага каждым из методов будет выполнена с использованием одних и тех же приближенно найденных величин.

Приведем общую постановку задачи оптимального управления. Пусть управляемый процесс описывается системой

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0, \quad t \in T = [t_0, t_1], \quad x(t) \in R^n, \quad u(t) \in R^r \tag{1}$$

с терминальными условиями

$$I_j(u) = \varphi^j(x, t) = 0, \tag{2}$$

и фазовыми ограничениями

$$J_i(u, t) = g^i(x, t) = 0, \quad i = \overline{1, s}, \quad t \in T. \tag{3}$$

Управление стеснено следующими ограничениями:

$$u(t) \in U, \tag{4}$$

где U – ограниченное замкнутое множество из R<sup>r</sup>. Вектор-функция f(x, u, t) непрерывно дифференцируема по x и u и непрерывна по t;  $\varphi^j(x)$ ,  $j = \overline{1, m}$ , – непрерывно дифференцируемые функции.

Пусть на правых концах траекторий системы (1) определена непрерывно дифференцируемая выпуклая скалярная функция

$$I_0(u) = \varphi^0(x(t_1)), \tag{5}$$

которую требуется минимизировать.

Градиенты функционалов  $I_j(u)$ ,  $j = \overline{0, m}$ , с помощью  $H^j(\psi_j, x, u, t) = \psi_j'(t) f(x, u, t)$  и сопряженной системы

$$\dot{\psi}_j = -f_x(x, u, t)' \psi_j(t), \quad \psi_j(t_1) = -\psi_j'(x(t_1)) \tag{6}$$

вычисляются по формулам

$$\nabla I_j(u) = -H_u^j(\psi_j, x, u, t), \quad j = \overline{1, m}. \tag{7}$$

Для каждого  $t \in T$  можно аналогично вычислить градиенты  $J_j(u, t)$ ,  $j = \overline{1, s}$ :

$$\nabla J_j(u, t) = -\overline{H}_u^j(\Phi_j, x, u, t, \tau), \quad t_0 \leq \tau \leq t \leq t_1, \tag{8}$$

где  $\overline{H}_u^j(\Phi_j, x, u, t, \tau) = \Phi_j'(t, \tau) f(x, u, \tau)$ ,  $\Phi_j(t, \tau)$ ,  $j = \overline{1, s}$ , – решения сопряженных систем  $\frac{\partial \Phi_j(t, \tau)}{\partial \tau} = -\frac{\partial f(x, u, \tau)}{\partial x} \Phi_j(t, \tau)$ ,  $\tau \in T$ ,  $\Phi_j(t, t) = -\frac{\partial g^j(x(t))}{\partial x}$ ,  $j = \overline{1, s}$ .

**1.2. Численные методы для решения задач с ограничениями на управление**

Перейдем теперь к рассмотрению алгоритмов для решения задач с ограничениями на управление (4), но без ограничений (2)-(3). Предположим, что при некотором  $u^k(t) \in U$ ,  $t \in T$ , найдено решение системы (1)  $x^k(t)$ ,  $t \in T$ .

Полагая  $j=0$ , проинтегрируем сопряженную систему от  $t = t_1$  до  $t = t_0$  при  $u = u^k(t)$ ,  $x = x^k(t)$ . На ее решении  $\psi^k = \psi^0(t)$  вычислим управление из принципа максимума:

$$\bar{u}^k(t) = \arg \max_{u \in U} H(\psi^k, x^k, u, t), \quad t \in T,$$

и найдем значения скалярной функции  $w_k(\bar{u}(t), t) = H(\psi^k, x^k, \bar{u}, t) - H(\psi^k, x^k, u^k, t)$ ,  $t \in T$ .

Если для заданного  $u^k$  и найденных  $x^k, \psi^k, \bar{u}^k$  принцип максимума нарушается:  $w_k(\bar{u}^k(\tau_k), \tau_k) > 0$ , то можно реализовать одну итерацию метода [3] для улучшения  $u^k$ .

Множество точек, в которых нарушается принцип максимума, обозначим через



$$T_\varepsilon = \{t \in T : w_k(\bar{u}^k(t), t) \geq \varepsilon w_k(\bar{u}^k(\tau_k), \tau_k)\}, \quad \varepsilon \in [0, 1],$$

где  $\tau_k$  – точка максимума этой функции на  $T_\varepsilon$ . Варьируя  $\varepsilon$ , можно найти его оптимальное значение  $\varepsilon_k$ , при котором управление

$$u_\varepsilon^{k+1}(t) = \begin{cases} \bar{u}^k(t), & t \in T_\varepsilon, \\ u^k(t), & t \in T \setminus T_\varepsilon, \end{cases} \quad (9)$$

доставит наименьшее значение целевому функционалу. При поиске  $\varepsilon_k$  можно использовать несколько потоков для одновременного интегрирования системы (1) с управлениями (9), соответствующими разным значениям  $\varepsilon_k$ .

В силу структуры управлений, генерируемых итерационной формулой (9), релаксации алгоритма могут прекратиться еще до получения управления, удовлетворяющего принципу максимума. Тогда для продолжения процесса оптимизации необходимо применить другой алгоритм, на итерациях которого конструируются управления не только с граничными, но и с внутренними относительно множества допустимых управлений значениями. Так, например, восстановить сходимость обычно удается с помощью построения выпуклой комбинации двух управлений:

$$u^{k+1}(t) = u^k(t) + \alpha[\bar{u}^k(t) - u^k(t)], \quad \alpha \in [0, 1]. \quad (10)$$

Вычисления по формулам (9) и (10) можно вести параллельно, выбирая из полученных приближений такое  $u^{k+1}$ , которому соответствует меньшее значение функционала. Если значения функционала будут сравниваться через несколько итераций, то в качестве критерия для сравнения эффективности методов (9) и (10) следует использовать значения приращений функционала, полученные на соседних итерациях каждого из методов. На практике установлено, что применение вариаций двух типов: "горизонтальной" (9) и "вертикальной" (10) позволяет избежать проявления эффекта "прилипания управления к границам", свойственного алгоритмам, основанным на принципе максимума.

Если в итерационной формуле (10) управление  $\bar{u}^k(t)$  будет вычисляться из линейризованного принципа максимума:

$$\bar{u}^k(t) = \arg \max_{u \in U} H_u(\psi^k, x^k, u, t)'u(t), \quad t \in T, \quad (10)$$

то получим итерации метода условного градиента. Очевидно, для систем, линейных по управлению, это управление будет совпадать с управлением, найденным из принципа максимума.

### 1.3. Численное решение задач с фазовыми ограничениями

Идея излагаемого ниже алгоритма [5] состоит в том, что на каждой его итерации решается вспомогательная задача минимизации модифицированной функции Лагранжа при линейризованных ограничениях (2), (3). Якобиан линейризованных ограничений строится из градиентов (7), (8), для расчета которых можно так же использовать параллельные вычисления. Значения двойственных переменных, полученные в результате решения вспомогательной задачи на каждой итерации, являются новым приближением для этих переменных на следующей итерации.

Пусть  $uk(t)$  – текущее приближение управления, а  $xk(t)$  – фазовая траектория, соответствующая  $uk(t)$ ,  $t \in T$ . Используя градиенты (7), (8), линейризуем условия (2), (3) в окрестности  $uk(t)$ :

$$I_i^L(u^k, u) = I_i(u^k) + \int_{T_0}^{T_1} \nabla I_i(u^k, t)'(u(t) - u^k(t))dt = 0 \quad i = \overline{1, m}, \quad (11)$$

$$J_j^L(u^k, u, \tau) = J_j(u^k, \tau) + \int_{t_0}^{\tau} \nabla J_j(u^k, t)'(u(t) - u^k(t))dt = 0, \quad (12)$$

$$j = \overline{1, s}, \quad \tau \in T$$

Построим модифицированную функцию Лагранжа для задачи (1)-(5) в следующем виде

$$L(u, u^k, \lambda^k, \mu^k) = I_0(u) - \lambda^{k'}(I(u) - I^L(u^k, u)) - \int_{t_0}^{t_1} \mu^k(t)'(J(u, t) - (13)$$

$$- J^L(u^k, u, t) dt + \frac{\rho}{2} (I(u) - I^L(u^k, u))' (I(u) - I^L(u^k, u)) + \\ + \frac{\rho}{2} \int_{t_0}^{t_1} (J(u, t) - J^L(u^k, u, t))' (J(u, t) - J^L(u^k, u, t)) dt,$$

где  $I, I^L$  –  $m$  – векторы,  $J, J^L$  –  $s$  – векторы,  $\lambda_k, \mu_k$  –  $m$  и  $s$  – мерные множители Лагранжа;  $\rho \geq 0$  – коэффициент штрафа.

На  $k+1$ -й итерации рассматриваемого метода решается задача минимизации функционала (13) на решениях системы (1) при линейных ограничениях (11), (12), (4). При этом для численного решения формируется задача математического программирования с линейными ограничениями специальной структуры.

## 2. Математическая модель деятельности предприятия

Сформулируем постановку задачи об оптимизации ставки налога на прибыль.

В нашем случае переменными состояниями будут затраты на производство  $x_1$  в момент  $t$ , налоговые отчисления  $x_2$  в момент  $t$ , количество основных производственных фондов  $x_3$  в момент  $t$ , количество трудовых ресурсов  $x_4$  в момент  $t$ . В качестве управления возьмем налоговую ставку  $v(t)$ .

Система уравнений, описывающая экономическую систему будет иметь вид [7]:

$$\frac{dx_1}{dt} = \mu \cdot (1 - v) \cdot F' - x_1;$$

$$\frac{dx_2}{dt} = \frac{1}{(1+r)^{t-t_0+1}} \cdot (F' - x_1)v;$$

$$\frac{dx_3}{dt} = b^{-1}(1-v)(1-\alpha)(1-u) \cdot F' - \chi x_3;$$

$$\frac{dx_4}{dt} = N' - x_4$$

Значения параметров данной задачи выбираются из интервалов, определяемых экономическим смыслом:

$0 < \mu < 1$  – доля выручки, которая составит затраты на производство в следующем периоде;

$0 < r < 1$  – рентабельность чистой прибыли;

$b \geq 0$  – коэффициент эффективности использования основных фондов;

$0 < \alpha < 1$  – коэффициент производственных затрат (норма материалоемкости);

$0 \leq u \leq 1$  – доля, которую составляет реальный темп непроизводственного потребления  $p(t)$

относительно его максимально возможной величины:  $u(t) = \frac{p(t)}{(1-\alpha(t))v(t)}$ ;

$\chi > 0$  – коэффициент амортизации;

$N'$  – функция, показывающая изменение численности персонала;

$F'$  – производственная функция.

Численность персонала определяется отношением затрат на оплату труда (затраты входят в затраты на производство) к средней заработной плате. Будем считать, что временной ряд для заработной платы имеет четко выраженную тенденцию, без циклической и сезонной составляющей, в нем не наблюдаются резкие темпы роста (снижения). Наиболее часто в таком случае при прогнозировании социально-экономических явлений, имеющих тенденцию, применяют полиномиальные функции. Примем так же, что процесс изменений заработной платы имеет равноускоренное изменение уровней ряда. Тогда функция изменения заработной платы будет иметь вид:

$$\tilde{Z}_t = n_1 \cdot t^2 + n_2 \cdot t + n_3,$$

где параметр  $n_1$  характеризует скорость изменения среднего прироста, параметр  $n_2$  – средний прирост, параметр  $n_3$  – уровень ряда при  $t=0$ . Следовательно, функция, характеризующая изме-

нения численности персонала принимает вид  $N' = \frac{f \cdot x_1}{n_1 t^2 + n_2 t + n_3}$ , где  $f$  – часть затрат на производство, которая составляет фонд оплаты труда.

В зависимости от характера производственного процесса, целей и средств моделирования в качестве производственных функций могут использоваться неотрицательные функции весьма разнообразного вида. Благодаря своей простоте и рациональности, широко применяется до сих пор и получила дальнейшие обобщения в различных направлениях производственная функция Кобба – Дугласа  $F' = A \cdot x_1^{a_1} x_3^{a_2} x_4^{a_3}$ , где  $a_{1,2,3}$  – показатели эластичности. Коэффициент  $A$ , стоящий перед знаком умножения, учитывает размерность, он зависит от выбранной единицы измерений затрат и выпуска. Обычно  $x_1$  и  $x_3$  измеряются в стоимостных единицах,  $x_4$  – в человеко-часах или численности работников. Такая производственная функция используется для описания объектов от промышленного объединения до отрасли, характеризующихся устойчивым, стабильным функционированием. Сумма коэффициентов эластичности представляет собой показатель эффекта расширения масштаба.

Как было сказано выше, производственная сфера функционирует при заданной налоговой ставке  $v(t)$  и своими внутренними управлениями она распоряжается так, чтобы минимизировать  $x_2$ , а внешняя среда заинтересована в том, чтобы налоговые отчисления были «побольше». В нашем случае задача оптимизации налоговой ставки состоит в максимизации налоговых отчислений  $x_2$ , то есть  $x_2 \rightarrow \max$ .

Формализованная постановка задачи об оптимизации ставки налога на прибыль будет иметь вид:

Переменные состояния:  $x_1, x_2, x_3, x_4$

Переменные управления:  $v(t)$

Уравнения процесса:

$$\frac{dx_1}{dt} = \mu \cdot (1 - v) \cdot (A \cdot x_1^{a_1} x_3^{a_2} x_4^{a_3}) - x_1;$$

$$\frac{dx_2}{dt} = \frac{1}{(1+r)^{t-t_0+1}} \cdot (A \cdot x_1^{a_1} x_3^{a_2} x_4^{a_3} - x_1)v;$$

$$\frac{dx_3}{dt} = b^{-1}(1-v)(1-\alpha)(1-u) \cdot A \cdot x_1^{a_1} x_3^{a_2} x_4^{a_3} - \chi x_3;$$

$$\frac{dx_4}{dt} = \frac{x_1 \cdot f}{n_1 t^2 + n_2 t + n_3} - x_4.$$

Начальные значения:

$$x_1(0) = x_{10}, x_2(0) = x_{20}, x_3(0) = x_{30}, x_4(0) = x_{40}, t \in [0, T]$$

Ограничения на фазовые переменные:

$$x_1 > 0, x_2 \geq 0, x_3 > 0, x_4 > 0$$

Ограничение на управление:

$$v_{\min} \leq v(t) \leq v_{\max}$$

Целевая функция:  $x_2 \rightarrow \max$

Параметры:

$$\mu, r, b, \alpha, u, \chi, a_0, a_1, a_2, a_3, f, d_1, d_2, d_3$$

Значения параметров являются постоянными:

$$\mu = \mu^*, r = r^*, b = b^*, \alpha = \alpha^*, u = u^*, v = v^*, \chi = \chi^*, a_0 = a_0^*, a_1 = a_1^*, a_2 = a_2^*, a_3 = a_3^*, f = f^*, d_1 = d_1^*, d_2 = d_2^*, d_3 = d_3^*.$$

### 3. Численные эксперименты

Проведем численное исследование модели на основе данных годовой отчетности 2002-2010 гг. ОАО «Иркутскэнерго» [8].

Доля выручки, которая составит затраты на производство в следующем году  $\mu$  определяется отношением затрат на производство к выручке (нетто) от продажи товаров за прошлый период:

$$\mu = \frac{36588536,00 \text{ т.р.}}{38838475,00 \text{ т.р.}} = 0,942069.$$

Рентабельность по чистой прибыли за 2010 год составила 18,1 %.

В качестве коэффициента эффективности использования основных фондов  $b$  используется фондоотдача, которая определяется как отношение прибыли до налогообложения к среднегодовой стоимости основных производственных фондов:

$$b = \frac{12063845,00 \text{ т.р.}}{35011802,50 \text{ т.р.}} = 0,345.$$

Коэффициент производственных затрат (норма материалоемкости)  $\alpha$  равен отношению затрат на производство к среднегодовой стоимости материальных ресурсов:

$$\alpha = \frac{2768012,50 \text{ т.р.}}{36588536,00 \text{ т.р.}} = 0,075652.$$

Разница между фактическим и чистым денежными потоками представляет собой непроизводственное потребление. Фактический денежный поток отражается в годовой отчетности по строке оборотные активы, а чистый денежный поток – это среднее арифметическое денежных средств на конец и начало года. Тогда

$$u = \frac{8986155,00 \text{ т.р.} - (234000,00 \text{ т.р.} - 271000,00 \text{ т.р.})}{(1 - 0,075652) \cdot 52352844,00 \text{ т.р.}} = 0,186459.$$

Коэффициент амортизации  $\chi$  равен отношению отчислений на амортизации к среднегодовой стоимости основных производственных фондов

$$\chi = \frac{2460000,00 \text{ т.р.}}{35011802,50 \text{ т.р.}} = 0,070262.$$

Ежегодно начиная с 2002 года затраты на заработную плату составляли в среднем 10% от затрат на производство. В это же время среднемесячная заработная плата увеличивалась. Точнее всего с коэффициентом детерминации  $R^2 = 99,5\%$  данные изменения характеризует полиномиальная функция с коэффициентами  $n_i > 0$ . Тогда

$$N^1 = \frac{f \cdot x_i}{n_1 t^2 + n_2 t + n_3}, \text{ где } n_1 = 224,75, n_2 = 1514,8, n_3 = 7111,3, f = 10\%.$$

Для производственной функции оценки параметров были получены методом наименьших квадратов  $A = 13,257, a_1 = 1,36, a_2 = 0,13, a_3 = 0,18$ . Сумма показателей эффективности больше единицы, то есть наблюдается возрастающая эффективность использования факторов производства (при одновременном возрастании на 1% всех факторов результат производства возрастет более чем на 1%).

Начальные значения для расчетов взяты из годовой отчетности:  $x_1(0) = 36588536$  т.р.,  $x_2(0) = 2631461$  т.р.,  $x_3(0) = 35011803$  т.р.,  $x_4(0) = 8072$  чел.

Для текущей ставки налога на прибыль сумма налоговых отчислений к концу исследуемого периода составит 14 166 908 тыс. руб.

Теперь найдем оптимальную ставку налога на прибыль для  $10\% \leq v(t) \leq 24\%$  и  $t \in [0,5]$  с помощью комплекса программ OPTCON [9]. Для каждого из приведенных в первом разделе класса задач в пакете OPTCON программно реализовано несколько методов оптимизации. При расчете оптимального управления в каждой конкретной задаче автоматически используется своя многомерная схема оптимизации. Рассмотрим динамику показателей деятельности ОАО «Иркутскэнерго» для полученной оптимальной ставки налога на прибыль (рис. 1):

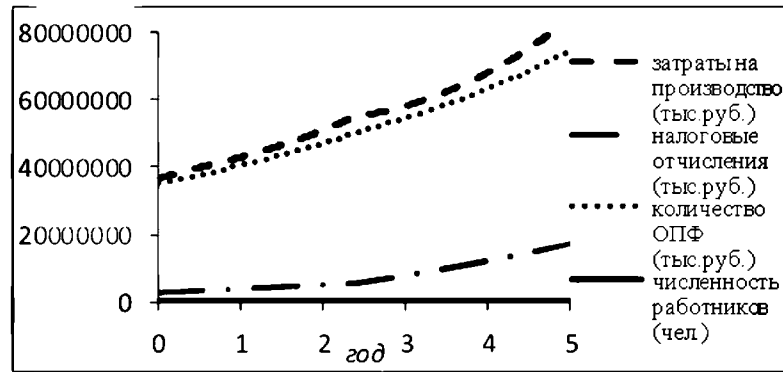


Рис. 1. Поведение модели при оптимальной ставке налога на прибыль

В ходе оптимизации ставки налогообложения прибыли предприятия было установлено, что на протяжении 5 лет наблюдается рост затрат на производство, налоговых отчислений, количества ОПФ и количества трудовых ресурсов. К окончанию рассматриваемого периода ставка налога на прибыль доходит до 24% (рис. 2), а налоговые отчисления достигают уровня 17,24 млрд р., вместо первоначальных 2,63 млрд р.

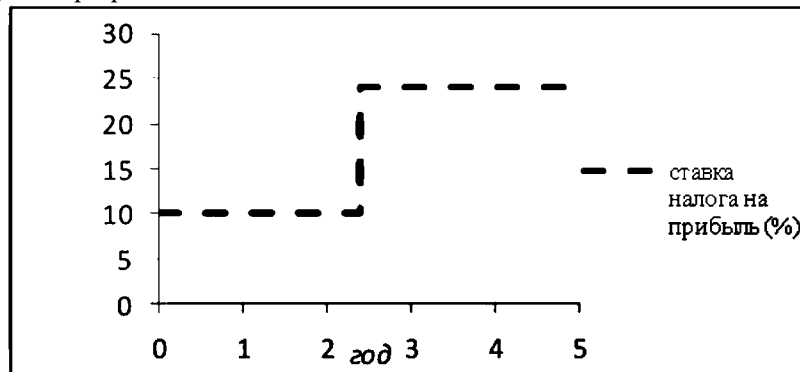


Рис. 2. Оптимальная ставка налога на прибыль предприятия

В обоих случаях наблюдается положительная динамика показателей деятельности предприятия. Но во втором случае сумма налоговых отчислений к концу периода на 12,1% больше по сравнению с налоговыми отчислениями при ставке налога на прибыль 20%. Но частая смена ставки может привести к непредсказуемым последствиям. Поэтому целесообразно установление ставки налога на прибыль в размере 15% для всего исследуемого периода. Тогда показатели деятельности для ОАО «Иркутскэнерго» будут следующими (рис. 3):

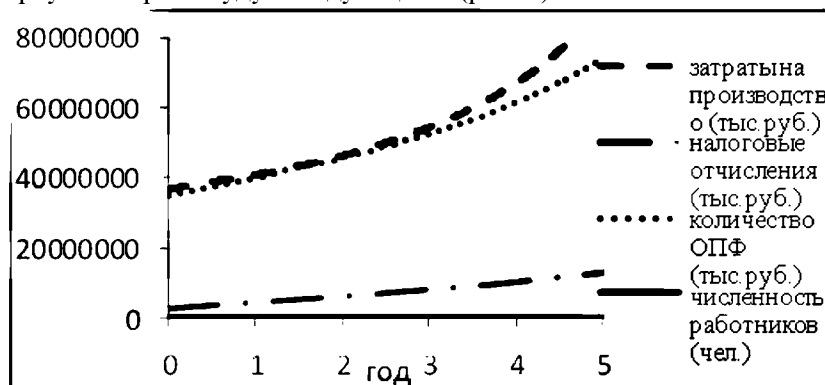


Рис. 3. Поведение модели при ставке налога на прибыль 15%

А сумма налоговых отчислений к концу периода (2015 г.) составит 12 931 644 тыс. р.

### Заключение

В ходе исследования и оптимизации ставки налогообложения прибыли предприятия были найдены значения параметров математической системы, моделирующей деятельность ОАО «Иркутскэнерго», ставки налога на прибыль и таких показателей, как затраты на производство, налоговые отчисления, количество основных производственных фондов, количество трудовых ресурсов. Оптимальная ставка налога на прибыль на основе проведенных расчетов составит 15%.

Если сравнивать ситуацию при начальных значениях и ситуацию, смоделированную при ставке налога на прибыль в 15%, то налоговые отчисления уменьшаются, т.е. выпадающие доходы по итогам пятилетнего периода составят 1 235 264 тыс. р. Но возместить их помогут инвестиционные проекты (доходы в виде инвестиций для последнего года исследуемого периода составят 8 819 010 тыс. р.). Такая ставка налога на прибыль позволит увеличить количество оборотных и инвестиционных средств в 1,5 раза к 2015 г.

Таким образом, можно сделать вывод, что для каждой задачи существует своя последовательность шагов, состоящая из итераций разных методов оптимального управления [3-6], которая адекватно учитывает особенности решаемой задачи и тем самым обеспечивает наиболее эффективный поиск оптимального управления. В многометодных алгоритмах построение такой последовательности выполняется автоматически по некоторому заданному критерию, оценивающему эффективность процесса оптимизации на всех этапах решения задачи.

На практике многометодная технология существенно повышает надежность получения правильных результатов при расчете оптимального управления реальными объектами, что в значительной мере обеспечивает эффективность и безопасность их движения по оптимальной траектории.

### Литература

1. Влияние налогообложения на инвестиции [Электронный ресурс]. URL: [www.investmarket.ru/Reports/show.asp?id=87](http://www.investmarket.ru/Reports/show.asp?id=87).
2. Налоговый кодекс РФ.
3. Васильев О.В., Тятюшкин А.И. Об одном методе решения задач оптимального управления, основанном на принципе максимума // Журн. вычисл. матем. и мат. физики. Т.21, №.6, 1981. С. 1376-1384.
4. Тятюшкин А.И. Параллельные вычисления в задачах оптимального управления // Сиб. журн. выч. матем. / РАН. Сиб. отд-ние. Новосибирск. 2000. Т. 3, № 2. С. 181-190.
5. Тятюшкин А.И. Численные методы решения задач оптимального управления с ограничениями на фазовые координаты // Изв. РАН. Теория и системы управления. 1998. № 2. С. 127-133.
6. Тятюшкин А.И. Многометодная технология оптимизации управляемых систем. Новосибирск: Наука, 2006. 343 с.
7. Москаленко А. И. Оптимальное управление моделями экономической динамики. Новосибирск: Наука, 1999.
8. Годовые отчеты ОАО «Иркутскэнерго» [Электронный ресурс]. URL: <http://www.irkutskenergo.ru/qa/768.2.html>.
9. Горнов А.Ю. Комплекс программ OPTCON для решения прикладных задач оптимального управления // Информационные и вычислительные технологии и системы: материалы всерос. конф. Улан-Удэ, 2003. Ч.1. С. 112-115.

*Колмакова Анастасия Ивановна*, аспирант Байкальского государственного университета экономики и права, [kolmakova.a.i@mail.ru](mailto:kolmakova.a.i@mail.ru).

*Kolmakova Anastasia Ivanovna*, postgraduate student from Baikal national university of economics and law.

## АЛГОРИТМ ГЛОБАЛЬНОГО ПОИСКА ГАРАНТИРОВАННЫХ РЕШЕНИЙ КВАДРАТИЧНО-ЛИНЕЙНОЙ ДВУХУРОВНЕВОЙ ЗАДАЧИ И ЕГО ТЕСТИРОВАНИЕ<sup>1</sup>

Рассматривается квадратично-линейная задача двухуровневой оптимизации с гарантированным (пессимистическим) решением. С использованием ее редукции к серии невыпуклых задач оптимизации предложены алгоритмы глобального и локального поиска. Приведены и проанализированы результаты численного решения случайно сгенерированных тестовых задач.

**Ключевые слова:** двухуровневые задачи оптимизации, гарантированное решение, локальный и глобальный поиск, генерация тестовых задач, вычислительный эксперимент.

*A. V. Malyshev*

## ALGORITHM OF GLOBAL SEARCH FOR GUARANTEED SOLUTIONS TO QUADRATIC-LINEAR BILEVEL PROBLEM AND ITS TESTING

A quadratic-linear bilevel optimization problem with guaranteed (pessimistic) solution is considered. Using its reduction to a series of nonconvex optimization problems global and local search algorithms are proposed. The results of computational solving randomly generated test problems are given and analyzed.

**Keywords:** bilevel optimization problems, guaranteed solution, local and global search, generating test problems, computational simulation.

### Введение

Задачи двухуровневой оптимизации представляют собой один из актуальных объектов исследования в современном математическом программировании [1–4]. Такие задачи зачастую возникают на практике при анализе (статических) иерархических систем управления в экономике, экологии и т.д. [2].

Для задач двухуровневой оптимизации не существует единого общепринятого понятия решения. Наиболее популярными являются оптимистическое и гарантированное решения. Подавляющее большинство работ, в которых исследуются задачи двухуровневой оптимизации, посвящено поиску в них оптимистических решений (например, обзор [5]). При этом нам удалось отыскать лишь несколько работ, посвященных поиску гарантированных решений в двухуровневых задачах [6, 7].

В настоящей статье развит новый подход к поиску гарантированных решений в двухуровневых задачах, в которых критерий эффективности игрока верхнего уровня – квадратичная функция, выпуклая по переменным игрока верхнего уровня и вогнутая по переменным игрока нижнего уровня, а критерий эффективности игрока нижнего уровня – линейная функция, ограничения на верхнем и нижнем уровнях иерархии также линейны. Этот подход основан на двухэтапной редукции [8, 9] рассматриваемой двухуровневой задачи к серии невыпуклых задач оптимизации и применении стратегии глобального поиска для численного решения последних [9, 10].

Статья организована следующим образом. В разд. 1 осуществляется постановка задачи и напоминает схема ее редукции. Далее, в разд. 2 предлагаются алгоритмы локального и глобального поиска в редуцированных задачах. Генерации тестовых двухуровневых задач с известными гарантированными решениями и численному их решению посвящен разд. 3.

### 1. Постановка задачи и ее редукция

В данной работе рассматривается задача поиска гарантированного решения квадратично-линейной двухуровневой задачи:

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ, проект №11-01-00270а.

$$\left. \begin{aligned} & \sup_y \{F(x, y) \mid y \in Y_*(x)\} \downarrow \min_x, \\ & x \in X, Y_*(x) = \mathop{\text{Arg min}}_y \{G(y) \mid y \in Y(x)\}, \end{aligned} \right\} \quad (\text{BP})$$

где  $F(x, y) = \frac{1}{2} \langle x, Cx \rangle + \langle c, x \rangle - \frac{1}{2} \langle y, C_1 y \rangle + \langle c_1, y \rangle$ ,  $G(y) = \langle d, y \rangle$ ,

$X = \{x \in R^m \mid Ax \leq a, x \geq 0\}$ ,  $Y(x) = \{y \in R^n \mid A_1 x + B_1 y \leq b, y \geq 0\}$ ,  $A \in R^{p \times m}$ ,  $A_1 \in R^{q \times m}$ ,  $B_1 \in R^{q \times n}$ ,  $C = C^T \geq 0, C_1 = C_1^T \geq 0$ .

В работе [8] при предположении ограниченности множеств  $X, Y(x)$  допустимых стратегий игроков осуществлено сведение задачи (BP) к серии невыпуклых задач оптимизации вида:

$$\left. \begin{aligned} & F(x, y) \downarrow \min_{x, y, v}, \\ & Ax \leq a, x \geq 0, A_1 x + B_1 y \leq b, y \geq 0, \\ & d - v c_1 + v C_1 y + v B_1 \geq 0, v \geq 0, \\ & h(x, y, v) = \langle d - v c_1 + v C_1 y, y \rangle + \langle b - A_1 x, v \rangle = 0, \end{aligned} \right\} \quad (\text{P}(v))$$

где  $v$  – вектор множителей Лагранжа для задачи нижнего уровня,  $v \geq 0$  – специальный параметр штрафа ( $v \downarrow 0$ ) [8].

На втором этапе редукции с использованием метода штрафов [2, 9] от задачи с нелинейным ограничением-равенством (P(v)) можно перейти к серии задач с выпуклым допустимым множеством и целевой d.c. функцией (представимой в виде разности двух выпуклых функций):

$$\left. \begin{aligned} & \Phi(x, y, v) = F(x, y) + \mu h(x, y, v) \downarrow \min_{x, y, v}, \\ & (x, y, v) \in D = \{(x, y, v) \mid Ax \leq a, x \geq 0, A_1 x + B_1 y \leq b, y \geq 0, \\ & d - v c_1 + v C_1 y + v B_1 \geq 0, v \geq 0\}, \end{aligned} \right\} \quad (\text{P}(\mu, v))$$

где  $\mu \geq 0$  – штрафной параметр.

В следующем разделе описаны алгоритмы локального и глобального поиска в задаче (P(μ, v)) при фиксированных значениях параметров штрафа  $\mu, v$ .

## 2. Локальный и глобальный поиск

Для осуществления локального поиска в задаче (P(μ, v)) при выборе значений параметров штрафа так, чтобы имело место неравенство

$$\mu v \geq \frac{1}{2}, \quad (1)$$

в работе [9] было предложено использовать минимизацию по отдельным группам переменных. Возникающие при этом вспомогательные оптимизационные задачи оказываются выпуклыми и, следовательно, могут быть эффективно решены с помощью классических численных методов [11]. Один из вариантов такого метода локального поиска (XY-процедуру) можно записать следующим образом.

Пусть задана точка  $(x_0, y_0, v_0) \in D$  и числовая последовательность  $\rho_s \downarrow 0$ .

**Шаг 0.** Положить  $s := 0, (x^s, y^s) := (x_0, y_0)$ .

**Шаг 1.** С помощью одного из методов линейного программирования найти приближенное решение  $v^{s+1}$  задачи:

$$\Phi(x^s, y^s, v) \downarrow \min_v, \quad (x^s, y^s, v) \in D,$$

с точностью  $\frac{\rho_s}{2}$ .

**Шаг 2.** Одним из методов квадратичного программирования найти приближенное решение  $(x^{s+1}, y^{s+1})$  задачи:

$$\Phi(x, y, v^{s+1}) \downarrow \min_{x, y}, \quad (x, y, v^{s+1}) \in D,$$



с точностью  $\frac{\rho_s}{2}$ .

**Шаг 3.** Положить  $s := s + 1$  и перейти на шаг 1.

Отметим, что для описанного выше варианта алгоритма локального поиска имеют место теорема сходимости к критической точке, являющейся (частично) глобальным решением задачи по переменным  $(x, y)$  и  $v$ , и критерии останова, аналогичные приведенным в [9] результатам для  $V$ -процедуры, в которой вспомогательные выпуклые задачи оптимизации решаются в другом порядке.

Далее, заметим, что целевую функцию задачи  $(P(\mu, \nu))$  можно представить в виде разности двух выпуклых функций, например, следующим образом:

$$\Phi(x, y, v) = g(x, y, v) - f(x, y, v), \quad (2)$$

$$g(x, y, v) = \left\langle x, \frac{1}{2}Cx + c \right\rangle + \frac{2\mu\nu - 1}{2} \langle y, C_1 y \rangle + \mu \left( \langle b, v \rangle + \frac{1}{4} \|v - A_1 x\|^2 \right) \text{ и}$$

$$f(x, y, v) = \langle (\mu\nu - 1)c_1 - \mu d, y \rangle + \frac{1}{4} \mu \|v + A_1 x\|^2 - \text{выпуклые функции (при выполнении неравенства (1)).}$$

Следовательно, задача  $(P(\mu, \nu))$  является задачей d.c. минимизации и для ее численного решения можно применить стратегию глобального поиска [9, 10], состоящую из двух основных этапов: а) локального поиска, позволяющего получить критическую точку; б) процедур выхода из текущей критической точки, основанных на условиях глобальной оптимальности. Алгоритм глобального поиска (АГП) в задаче  $(P(\mu, \nu))$ , основанный на данной стратегии, имеет следующий вид.

Пусть дана некоторая точка  $(x_0, y_0, v_0) \in R^{m+n+q}$ , числовые последовательности  $\tau_k \downarrow 0$ ,  $\delta_k \downarrow 0$ , множество направлений  $Dir$  ( $N$  – количество направлений в данном множестве), числа  $\gamma_- = \inf(g, D)$  и  $\gamma_+ = \sup(g, D)$  и параметры алгоритма  $M, \xi$ .

**Шаг 0.** Положить  $k := 0, l := 1, \gamma := \gamma_-, \Delta\gamma := (\gamma_+ - \gamma_-) / M$ .

**Шаг 1.** Построить точку  $(\tilde{x}^k, \tilde{y}^k, \tilde{v}^k) \in D$  как приближенное решение задачи

$$\|(x - x_k, y - y_k, v - v_k)\|^2 \downarrow \min_{x, y, v}, \quad (x, y, v) \in D,$$

с точностью  $\tau_k$  одним из методов квадратичного программирования.

**Шаг 2.** Начиная из точки  $(\tilde{x}^k, \tilde{y}^k, \tilde{v}^k)$ , с помощью  $XY$ - или  $V$ -процедуры построить приближенно критическую точку  $(x^k, y^k, v^k) \in D$  с точностью  $\tau_k$ ; положить  $\zeta_k := \Phi(x^k, y^k, v^k)$ .

**Шаг 3.** По точке  $(\tilde{z}^l, \tilde{u}^l, \tilde{w}^l) \in Dir$  построить точку  $(z^l, u^l, w^l) = \lambda_{l,k}(\tilde{z}^l, \tilde{u}^l, \tilde{w}^l)$  из аппроксимации  $A_k$  поверхности уровня функции  $f(\cdot): f(z^l, u^l, w^l) = \gamma - \zeta_k$ .

**Шаг 4.** Если  $g(z^l, u^l, w^l) > \gamma + \xi\gamma$ , то положить  $l := l + 1$  и перейти на шаг 3.

**Шаг 5.** Построить точку  $(z^l, u^l, w^l) \in D$  как приближенное решение частично линейаризованной задачи

$$g(x, y, v) - \left\langle \nabla f(\tilde{z}^l, \tilde{u}^l, \tilde{w}^l), (x, y, v) \right\rangle \downarrow \min_{x, y, v}, \quad (x, y, v) \in D,$$

с точностью  $\delta_k$  одним из методов квадратичного программирования.

**Шаг 6.** Начиная с точки  $(z^l, u^l, w^l)$ , с помощью  $XY$ - или  $V$ -процедуры построить приближенно критическую точку  $(\tilde{x}^l, \tilde{y}^l, \tilde{v}^l) \in D$  с точностью  $\delta_k$ .

**Шаг 7.** Если  $\Phi(\tilde{x}^l, \tilde{y}^l, \tilde{v}^l) < \Phi(x^k, y^k, v^k)$ , то положить  $(\tilde{x}^{k+1}, \tilde{y}^{k+1}, \tilde{v}^{k+1}) := (\tilde{x}^l, \tilde{y}^l, \tilde{v}^l)$ ,  $k := k + 1, l := 1, \gamma := \gamma_-$  и перейти на шаг 2.

**Шаг 8.** Если  $\Phi(\tilde{x}^l, \tilde{y}^l, \tilde{v}^l) \geq \Phi(x^k, y^k, v^k)$  и  $l < N$ , то положить  $l := l + 1$  и вернуться на шаг 3.

**Шаг 9.** Если  $\Phi(\tilde{x}^l, \tilde{y}^l, \tilde{v}^l) \geq \Phi(x^k, y^k, v^k)$ ,  $l = N$  и  $\gamma < \gamma_+$ , то положить  $\gamma := \gamma + \Delta\gamma, l := 1$  и вернуться на шаг 3.

**Шаг 10.** Если  $\Phi(\tilde{x}^l, \tilde{y}^l, \tilde{v}^l) \geq \Phi(x^k, y^k, v^k)$ ,  $l = N$  и  $\gamma = \gamma_+$ , то стоп;  $(x^k, y^k, v^k)$  – полученное решение задачи  $(P(\mu, v))$ .

Отметим, что на шагах 0 и 1 АПП осуществляется инициализация алгоритма и построение допустимой начальной точки, на шаге 2 – локальный поиск. Шаги 3-10 АПП соответствуют процедурам выхода из текущей критической точки, основанным на условиях глобальной оптимальности [9, 10]. Условием выхода из текущей критической точки является улучшение по целевой функции, проверяемое на шаге 7.

Отдельно отметим, что в настоящей работе один из ключевых моментов глобального поиска – выбор множества направлений  $Dir$  – осуществлялся точно так же, как и в [12], с использованием текущей критической точки и векторов евклидового базиса. При этом были выбраны следующие значения параметров  $M, \xi$ , отвечающих за точность и скорость работы алгоритма:  $M = 5, \xi = 0$ .

В следующем разделе перейдем к численному тестированию описанного алгоритма глобального поиска.

### 3. Тестирование алгоритма глобального поиска

Одним из важных вопросов при тестировании новых численных методов является выбор поля тестовых примеров. В данной работе с этой целью был предложен метод генерации тестовых задач, основанный на методике из [13], базирующийся на построении тестовых задач из задач-ядер небольшой размерности. В частности, были составлены и решены аналитически задачи-ядра следующего вида ( $m = 1, n = 2$ ):

$$\left. \begin{aligned} & \sup_y \{x^2 - 8x + py_1 - 2y_2^2 \mid y \in Y_*(x)\} \downarrow \min_x, \quad x \in [0; 6], \\ & Y_*(x) = \underset{y}{\text{Arg min}} \{-y_1 \mid y_1 + y_2 \leq x, 3 \geq y_1 \geq 0, y_2 \geq 0\}, \end{aligned} \right\} \quad (\text{KP})$$

где  $p \in \{3, 4, 6\}$  – параметр, влияющий на структуру локальных и глобальных решений задачи.

**Предложение.** При  $p = 3$  задача (KP) имеет локальное решение в точке  $x = 2.5$  и глобальное решение в точке  $x = 4$ . При  $p = 4$  задача (KP) имеет два глобальных решения в точках  $x = 2, 4$  и не имеет локальных решений, не являющихся глобальными. В случае  $p = 6$  задача (KP) имеет глобальное решение в точке  $x = 1$  и локальное решение в точке  $x = 4$ .

Далее, согласно методике из [13] осуществляется построение задач произвольной размерности  $m = r, n = 2r$  объединением задач-ядер вида (KP) при различных значениях параметра  $p$ . Наконец, осуществляется невырожденное линейное преобразование координат, полученное с помощью генератора псевдослучайных чисел, чтобы избавиться от сепарабельности построенной таким образом задачи. Количество локальных решений, не являющихся глобальными, при этом известно и может достигать астрономических величин.

При проведении вычислительного эксперимента вышеописанным методом были сгенерированы серии двухуровневых задач размерности от  $m = 5, n = 10$  до  $m = 35, n = 70$ . На полученном поле тестовых примеров был протестирован алгоритм глобального поиска из разд. 2, запрограммированный на C++. При этом для решения вспомогательных выпуклых задач оптимизации использовался пакет FICO Xpress Optimization Suite, который считается специалистами одним из лучших средств в данной области. Использовался компьютер с процессором Intel Core 2 Duo 2.0GHz. Остальные условия вычислительного эксперимента по сравнению с [9] не изменились.

В таблице 1, где приведены результаты вычислительного эксперимента, использовались следующие обозначения:

$m + n$  – суммарная размерность (по переменным  $x$  и  $y$ ) каждой тестовой двухуровневой задачи в серии тестовых задач;

$N$  – количество задач в серии;

LocSol – среднее количество локальных решений, не являющихся глобальными, в задачах серии;

Loc – среднее число запусков алгоритма локального поиска при проведении глобального поиска в серии задач;

St – среднее количество итераций алгоритма глобального поиска;

T – среднее время работы программы, реализующей алгоритм глобального поиска.

Прежде всего, отметим, что во всех сгенерированных тестовых задачах с помощью алгоритма глобального поиска удалось найти глобальные решения с точностью  $\varepsilon = 10^{-3}$ . При этом количество пройденных алгоритмом глобального поиска критических точек, позволивших улучшить значение целевой функции, невелико и достигло в среднем лишь 13.6 для задач размерности 105. Также среднее количество запусков процедуры локального поиска, достигающее 5863.4, несравнимо с огромным количеством локальных решений, не являющихся глобальными, достигающим значения  $3.4 \cdot 10^{10}$ .

Таблица 1

Тестирование глобального поиска на сериях задач

$m+n$	N	LocSol	Loc	St	T
15	10	28.0	30.5	1.5	1:29
30	10	951.5	150.1	1.5	17:17
45	10	32680.0	447.7	2.7	1:43:16
60	10	932783.0	746.2	4.9	5:08:00
75	10	$3.3 \cdot 10^7$	799.1	4.1	8:01:57
90	10	$1.0 \cdot 10^9$	1462.9	5.9	20:46:09
105	10	$3.4 \cdot 10^{10}$	5863.4	13.6	1:51:21:80

В целом, по результатам вычислительного эксперимента заключаем, что использование более эффективного пакета для решения вспомогательных выпуклых задач оптимизации, возникающих при глобальном поиске, позволило существенно повысить размерность решаемых задач (по сравнению с [9]).

### Заключение

В работе предложены алгоритмы локального и глобального поиска в квадратично-линейных двухуровневых задачах с гарантированным решением. В результате вычислительного эксперимента алгоритмом глобального поиска за приемлемое время были решены все сгенерированные тестовые задачи до рекордной (для данного класса задач) размерности 105.

### Литература

1. Гермейер Ю.Б. Игры с противоположными интересами. – М.: Наука, 1976.
2. Bard J.F. Practical Bilevel Optimization. – Dordrecht, The Netherlands: Kluwer Academic Publishers, 1998.
3. Dempe S. Foundations of Bilevel Programming. – Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.
4. Pang J.-S. Three modeling paradigms in mathematical programming // Mathematical Programming. 2010. Vol. 124, № 2. P. 297-323.
5. Colson B., Marcotte P., Savard G. An overview of bilevel optimization // Annals of operations research. 2007. Vol. 153, № 1. P. 235-256.
6. Молодцов Д.А. О решении одного класса неантагонистических игр // Журн. вычисл. матем. и матем. физики. 1976. Т. 16, № 6. С. 1451-1456.
7. Tsoukalas A., Wiesemann W., Rustem B. Global Optimisation of Pessimistic Bi-Level Problems // Lectures on global optimization / Ed. by P.M. Pardalos, T.F. Coleman. Toronto, Canada. 2009. Vol. 55. P. 215-243.
8. Мальшев А.В., Стрекаловский А.С. О взаимосвязи некоторых задач двухуровневой и нелинейной оптимизации // Известия вузов. Математика. 2011. № 4. С. 99-103.
9. Мальшев А.В., Стрекаловский А.С. Глобальный поиск гарантированных решений в квадратично-линейных задачах двухуровневой оптимизации // Известия Иркутского государственного университета. Математика. 2011. Т. 4, № 1. С. 73-82.
10. Стрекаловский А.С. Элементы невыпуклой оптимизации. – Новосибирск: Наука, 2003.
11. Васильев Ф.П. Методы оптимизации. – М.: Факториал-пресс, 2002.
12. Стрекаловский А.С., Орлов А.В., Мальшев А.В. Численное решение одного класса задач двухуровневого программирования // Сибирский журнал вычислительной математики. 2010. Т. 13, № 2. С. 201-212.
13. Calamai P., Vicente L. Generating Linear and Linear-Quadratic Bilevel Programming Problems // SIAM Journal on Scientific Computing. 1993. Vol. 14, № 4. P. 770-782.

*Мальшев Антон Валентинович*, канд. физ.-мат. наук, программист, Институт динамики систем и теории управления СО РАН, 664033, Иркутск, ул. Лермонтова, 134, тел. (395-2) 453082, e-mail: [anton@irk.ru](mailto:anton@irk.ru).

*Malyshev Anton Valentinovich*, candidate of physical and mathematical sciences, programmer, Institute for system dynamics and control theory SB RAS.

## 2. Системный анализ и информационные технологии

УДК 004.4'244

© А.А. Ветров, Е.С. Фереферов, А.Е. Хмельнов, А.О. Шигаров

### ФОРМИРОВАНИЕ ХРАНИЛИЩА ДАННЫХ ДЛЯ СИСТЕМЫ MDATTR<sup>1</sup>

Статья посвящена актуальным задачам формирования хранилища данных и конфигурированию метаописаний, для их совместного использования в информационной системе MDAttr.

**Ключевые слова:** хранилище данных, анализ данных, интеграция данных.

*A.A. Vetrov, E.S. Fereferov, A.E. Hmelnov, A.O. Shigarov*

### FORMING DATA WAREHOUSE FOR SYSTEM MDATTR

The article is devoted to actual problems of creating a data warehouse and forming metadescriptions for their use in MDAttr system.

**Keywords:** data warehouse, data analysis, data integration.

#### Введение

Принятие управленческих решений в крупных организациях базируются на комплексном анализе распределенных и разноформатных баз данных (БД) большого объема. Это обуславливает актуальность их интеграции путем формирования хранилища данных (ХД) [1-2]. Создание ХД позволяет проанализировать разрозненные данные из корпоративной сети, находящиеся в форматах различных систем управления базами данных (СУБД), и обеспечить выполнение сложных аналитических запросов к информации.

#### 1. Система MDAttr

Система MDAttr (рис.1), созданная в ИДСТУ СО РАН, предназначена для анализа многомерных числовых данных, зависящих от различных характеристик. Система разработана с использованием технологии метаописаний (спецификаций) баз данных (БД) [3], которая активно используется для разработки настольных приложений, работающих с БД [4-5].

		2000г.	2001г.	2002г.	2003г.	2004г.
Энгельсов район	Трудовые ресурсы, Тысяча человек		7,7	7,8		
	Всего земель, Гектар		1132370			
Иркутский район	Трудовые ресурсы, Тысяча человек		31,396	32,279		
	Всего земель, Гектар		3327578			
Казачинско-Ленский район	Трудовые ресурсы, Тысяча человек		13,2	13,396		
	Всего земель, Гектар		13904330			
Катангский район	Трудовые ресурсы, Тысяча человек		2,996	2,771		
	Всего земель, Гектар		3140884			
Качугский район	Всего земель, Гектар					

Рис. 1. Интерфейс MDAttr

<sup>1</sup> Работа выполнена при финансовой поддержке грантов РФФИ 11-07-92204-Монг\_а, 11-07-00426-а, проекта ОНИТ СО РАН №41 и междисциплинарного проекта СО РАН № 131.

Система реализует:

- многомерное концептуальное представление данных с полной поддержкой множественных иерархий;
- технологию OLAP-анализа;
- генерируемый графический пользовательский интерфейс на основе метаописаний;
- поддержку большинства современных систем управления базами данных (СУБД) через интерфейс ADO (ActiveX Data Objects) [6];
- привязку пространственных данных с возможностью построения тематических карт;
- вычисляемые поля, задаваемые выражениями;
- многоязычный пользовательский интерфейс;
- построение диаграмм по результатам анализа;
- экспорт данных в электронные таблицы, для использования сторонними системами.

ХД выступает источником данных и метаописаний для MDAAttr, и настраивается администратором под конкретную предметную область, причём конфигурация определяется требованиями MDAAttr.

## 2. Технология метаописаний баз данных

Данная технология выводит за пределы конкретной информационной системы (ИС) общую функциональность, создавая ядро ИС, которое используется в разрабатываемых системах, при этом информация о предметной области хранится в метаописаниях и используется ядром для построения динамически генерируемого пользовательского интерфейса приложения. Технология позволяет абстрагироваться от структуры БД и используемой СУБД за счет представления в метаописании информации о структуре БД, привязке к карте, пользовательском интерфейсе, и реализации динамической генерации интерфейса приложения на основе такой информации. Метаописания хранятся в структурированном машиночитаемом виде в текстовых файлах или в таблицах ХД на сервере [7]. Для формирования спецификаций разработан язык представления баз данных [8], позволяющий в декларативном виде описывать как структуру данных (способ подключения к БД, структуры таблиц БД, связей между ними), так и структуру приложения (представления таблиц, способы отображения данных), а также способы интеграции с другими БД и ГИС.

Использование данной технологии снижает сложность разработки и позволяет сократить сроки создания реальных информационных систем в сравнении с традиционными подходами, связанными с описанием объектов и связей между ними средствами визуального программирования и в коде приложения.

## 3. Конфигурирование метаописаний хранилища данных

Для существующей версии системы MDAAttr таблицы метаописаний в ХД заполняются с использованием запросов на языке SQL. Рассмотрим процесс создания метаописаний на примере простого ХД. Пусть в ХД находится информации о населении. Данные собраны по отраслям деятельности, территориям и времени (три таблицы измерений). На рис. 2 представлены две модели данных рассматриваемого примера: модель рассматриваемой предметной области и модель, используемая в ИС MDAAttr.

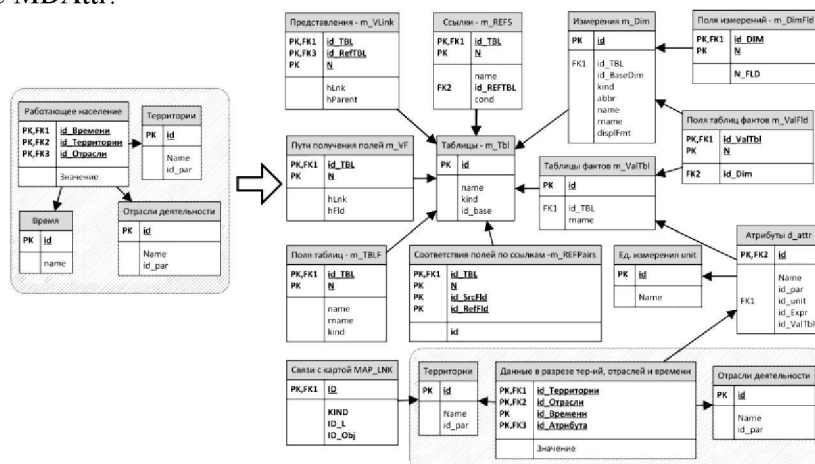


Рис. 2. Пример ХД

Данные таблицы «Население» зависят от трёх измерений: территории, отрасли, время. Связь таблицы с атрибутом «население» заносится в таблицу «Атрибуты». В эту же таблицу можно поместить и другие данные собираемые в данном разрезе (территории, отрасли, время), для этого требуется добавить поле с кодом атрибута. Значения атрибутов, собираемых в одном разрезе, удобно хранить в одной таблице, а для разрезов с разными количествами влияющих измерений необходимо использовать разные таблицы значений атрибутов.

Для подготовки метаописаний необходимо выполнить следующий алгоритм:

- Создать 13 таблиц с predetermined структурой, в которых будет содержаться вся метаинформация.
  1. m\_tbl - информация о перечне всех таблиц
  2. m\_tblf - информация о полях таблиц из m\_Tbl
  3. m\_vLink - информация о представлениях
  4. m\_vf - информация о путях получения полей m\_tblf для m\_VLink
  5. m\_refs - информация о связях (ссылках) между таблицами
  6. m\_refspairs - пары полей, связывающие таблицы из m\_refs
  7. m\_dim - информация обо всех измерениях хранилища, хранящихся в перечне таблиц m\_tbl
  8. m\_dimfld - описание используемых полей составляющих измерения m\_dim
  9. m\_valtbl - информация о таблицах из m\_tbl, которые хранят факты
  10. m\_valfld - информация о полях таблиц из m\_valtbl со ссылками на измерения m\_dim
  11. d\_attr - информация о показателях
  12. unit - единицы измерения
  13. map\_lnk - связь данных измерений территорий с объектами на карте
- Заполнить перечень таблиц ХД, записав его в m\_tbl и указав тип таблицы (справочник, простая таблица, представление).
- Заполнить список полей каждой таблицы, записав его в m\_tblf, ссылаясь на таблицы и указав тип (строка, число, ссылка) и имя, которое будет использоваться в интерфейсе пользователя.
- Заполнить таблицу m\_VLink, в соответствии с записями о представлениях из m\_tbl, указав для каждого представления, из каких таблиц оно создается.
- Заполнить m\_VF информацией, из каких полей каких таблиц создается представление, ссылаясь на таблицу m\_VLink.
- Заполнить таблицу m\_refs, указав какие таблицы с какими связаны. Например, нужно создать три записи для таблицы Данные (рис. 3), которая ссылается на три таблицы: Время, Отрасли, Территории.
- Заполнить таблицу m\_refspairs, указав, по каким полям связаны таблицы между собой таблицы. Например, таблица Данные (id=6) своим полем id=1 ссылается на таблицу Территории (id=3) на её первичный ключ (id=0).
- Заполнить таблицу m\_Dim информацией из m\_tbl указав какие таблицы являются измерениями, отметив отдельно измерения территорий и временем. Здесь присваиваются названия измерениям, которые будут использоваться для построения пользовательского интерфейса.

id	name	kind	id_base							
1	0	Ед.измерения	L	0						
2	1	Время	L	0						
3	2	Атрибуты	T	0						
4	3	Территории	T	0						
5	4	Отрасли	T	0						
6	5	m_ValTbl	T	0	23	4	0	id	№	I
7	6	Данные	T	5	24	4	1	name	Отрасль	N
8	7	vAttr Дерево атрибутов	T	2	25	4	2	id_par	№ родителя	~

Рис. 3. Пример таблицы m\_tbl и фрагмента таблицы m\_tblf

- Заполнить таблицу m\_DimFld информацией, какие поля из измерений в m\_dim будут использоваться в информационной системе.
- Заполнить таблицу m\_ValTbl информацией о таблицах фактов.
- Заполнить таблицу m\_ValFld информацией, какие поля относятся к каким измерениям, и какое поле является фактом (числовым значением).

id	id_TBL	kind	abbr	name	id_ValTbl	N	id_Dim		
1	0	3	T	O	Территории	1	1	0	2
2	1	4	T	D	Отрасли	2	1	1	0
3	2	-1	D	T	Периоды (Время)	3	1	2	1
4	3	7	T	A	Показатели	4	1	3	-1

id	id_TBL	name	
1	1	9	Данные по работающему населению

Рис. 4. Содержимое таблиц m\_Dim (слева), m\_ValTbl (снизу) и m\_ValFld (справа).

- Если необходимо, заполнить таблицу map\_Ink числовыми параметрами-ссылками. Связи между объектами карты и измерением территорий.

После выполнения всех этапов алгоритма, ХД считается заполненным метаописаниями. На рассматриваемом примере видно, что заполнение метаописаний ХД является нетривиальной работой, так как трудно проследить связи между записями таблиц и безошибочно провести заполнение. В реальных системах, в отличие от рассмотренного примера, количество измерений и фактов на порядок больше, что еще усложняет процесс конфигурирования.

#### 4. Наполнение хранилища данных

Часто исходная информация, необходимая для формирования ХД, доступна только в неструктурированном или слабоструктурированном виде, например, в виде статистических или финансовых отчетов. Одним из наиболее используемых способов представления такой информации являются содержащиеся в текстах таблицы. Обычно такие таблицы не содержат полного набора структурных метаданных, необходимых для машинной обработки табличной информации, в том числе, для загрузки в ХД. Даже высокоуровневое представление таблицы, например, в виде HTML разметки или объекта текстового процессора Word, не включает информацию о разделении ячеек на данные и заголовки, связях между ячейками, а также об используемых типах данных. Извлечение информации из таких таблиц требует восстановления отсутствующих метаданных об их структуре. Для решения этой задачи авторами развиваются оригинальные методы и программное обеспечение [9-14].

В частности, для автоматизации массового ввода информации из слабоструктурированных таблиц со сложной компоновкой, представленных в формате табличного процессора Excel, в реляционные базы данных разработана программная система Excel2DB [11]. В отличие от различных ETL средств, которые позволяют рассматривать файлы табличного процессора в качестве источника данных, система Excel2DB позволяет загружать данные из таблиц, имеющих более сложную структуру, чем просто «решеточную», при которой строки обрабатываемой таблицы в формате Excel непосредственно соответствуют записям целевой таблицы БД, а столбцы – полям. Для анализа и преобразования табличной информации в программе Excel2DB используются определяемые пользователями модели таблиц. В Excel2DB модель позволяет в декларативной манере задать физическую и частично логическую компоновку таблицы в пространстве ячеек: границы таблицы внутри листа; позиции областей ячеек, содержащих атрибуты (заголовки), и областей, содержащих факты (данные); областей, физическая компоновка которых задается не структурой ячеек листа, а разграфкой; уровни атрибутов и подчиненность между ними; тиражируемость фактов в столбце и др. Такие модели достаточно строго задают вид обрабатываемых таблиц. Поэтому предлагаемая технология эффективна (требует меньших затрат) в тех случаях, когда заданная модель описывает некоторый набор таблиц. Т.е. извлечение, преобразование и загрузку информации из такого набора в базу данных можно выполнить в Excel2DB с помощью одной модели. За счет использования настраиваемых моделей Excel2DB позволяет обрабатывать широкий класс таблиц. Однако сами модели ограничены рядом предположений о таблицах, в терминах которых они задаются. Расширение возможностей определения моделей возможно только путем изменения алгоритмов и структур данных системы Excel2DB.

Другим примером автоматизации процессов извлечения и преобразования информации из таблиц, содержащихся в текстах, является разработанная авторами система TabExtractor [12]. Система основана на использовании алгоритмов со встроенными правилами анализа и обработки таблиц, ориентированными на особенности компоновки таблиц из государственных и корпоративных статистических отчетов, в частности, отчетов, публикуемых Росстатом. Поэтому по сравнению с Excel2DB система TabExtractor позволяет обрабатывать более узкий класс таблиц, однако при



этом в процессе обработки табличной информации не требуется задавать модели таблиц. Входными данными системы являются файлы печати в формате Enhanced Metafile. Документы различных форматов (Word, Excel, ASCII текст, HTML и др.) могут быть напечатаны в такие файлы печати. В системе TabExtractor процесс извлечения информации из файлов печати включает следующие этапы: 1) извлечение текста, текстовых метрик и разграфки из файлов печати; 2) анализ физической компоновки таблицы, включая обнаружение и сегментацию таблиц; 3) анализ логической компоновки, включая разделение табличного содержания на атрибуты и факты, а также восстановление связей между ними. При необходимости описанный процесс может включать ручную обработку. В результате табличная информация преобразуется к структурированному виду, для которого остается выполнить восстановление связей атрибутов с измерениями в терминах реляционной модели данных, чтобы выполнить ее загрузку в реляционную базу данных. Этот процесс может выполняться вручную или полуавтоматически в программах AttrTree [13] и СтатИмп [14].

### 5. Средство конфигурирования метаописаний хранилища данных

Авторами разрабатывается инструментальное средство Администратор MDAttr, позволяющее неподготовленному специалисту провести конфигурирование ХД для системы MDAttr с помощью интуитивно понятного графического интерфейса вместо SQL запросов.

Рассматриваемое средство представляет собой визуальный пользовательский интерфейс, автоматизирующий все этапы рассмотренного алгоритма формирования метаописаний, решает задачи создания и удаления таблиц, измерений и фактов ХД.

Данное средство позволяет:

- настраивать соединение с хранилищем путем построения строки подключения и подключаться к ХД системы MDAttr;
- считывать и записывать метаописания;
- управлять описанием таблиц, представлений, измерений и фактов;
- настраивать список таблиц, задействованных в системе MDAttr
- просматривать содержимое таблиц;
- обеспечивает ссылочную целостность на уровне метаописаний;
- предоставляет интерфейс привязки объектов ХД к карте, путем заполнения пар значений код объекта на карте / код объекта в ХД.

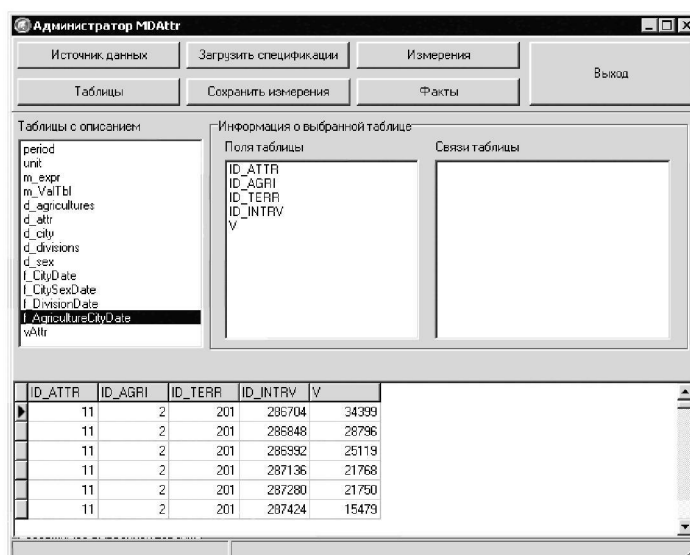


Рис. 6. Главная форма средства администрирования

Логически система делится на три модуля для работы с таблицами, измерениями и фактами соответственно.

Модуль работы с таблицами позволяет проводить наиболее общие операции с метаописаниями:

- создавать и удалять набор таблиц для хранения метаописаний;



- управлять отображением полей таблиц и редактировать их названия для отображения в интерфейсе MDAtr;
- конфигурировать связи между таблицами;
- определять представления.

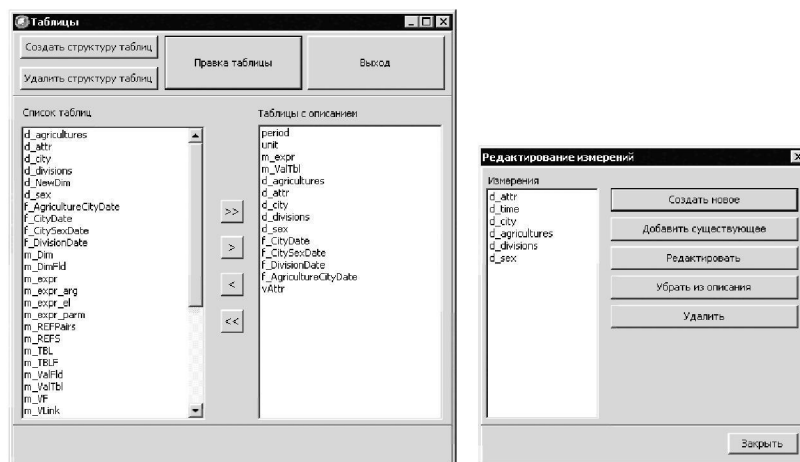


Рис. 6. Формы для работы с таблицами и измерениями ХД

Модуль измерений и модуль фактов похожи по своей функциональности и предоставляют возможности:

- создавать новые измерения и факты, указывая название и тип полей нового измерения;
- редактировать метаописания существующих измерений и фактов ХД;
- добавлять метаинформацию к существующим измерениям в ХД (у которых еще нет метаописания);
- редактировать метаинформацию имеющихся измерений;
- простым выбором указывать таблицу, содержащую факты ХД;
- исключать измерения и факты из метаописаний;
- полностью удалять измерения и факты из ХД.

### Заключение

Рассматриваемое инструментальное средство решает важные задачи конфигурирования метаописаний ХД, позволяя через визуальный интерфейс автоматизировать приведенный алгоритм формирования спецификаций ХД, при этом снижается его сложность и упрощается управление метаописаниями. Администратор MDAtr используется в совместном российско-монгольском проекте, посвященном анализу показателей социально-экономического развития территорий республики Монголия, направленному на повышение эффективности работы территориальных органов управления.

### Литература

1. Inmon W. Building the Data Warehouse. — New York: Wiley, 2005. — 576 с.
2. Kimball R., Rose M. The Data Warehouse Toolkit. — New York: Wiley, 2002. — 464 с.
3. Бычков И.В., Фереферов Е.С., Хмельнов А.Е. Метаописание баз данных как основа интеграции информационно-справочных систем и ГИС // Вычислительные технологии. 2007. Т. 12, № 5. С. 41-51.
4. Внедрение современных информационных технологий в региональных проектах / И.В. Бычков [ и др.] // Вестник НГУ. Серия: Информационные технологии. 2008. Т. 6, № 1. С. 15-24
5. Муниципальная информационная система обеспечения градостроительной деятельности / А.С. Гаченко [и др.] // Вычислительные технологии. 2008. Т. 13, № 1. С. 11-16.
6. Microsoft ActiveX Data Objects (ADO), 2012. URL: [http://msdn.microsoft.com/en-us/library/windows/desktop/ms675532\(v=vs.85\).aspx](http://msdn.microsoft.com/en-us/library/windows/desktop/ms675532(v=vs.85).aspx) (дата обращения май 2012 года)
7. Ветров А.А., Фереферов Е.С., Хмельнов А.Е. Технология использования метаописаний для формирования хранилищ данных и анализа многомерных данных // Вестник Бурятского государственного университета. 2011. № 9. С. 83-86.

8. Фереферов Е. С., Хмельнов А. Е. Язык представления баз данных // Материалы III Международной конференции Инфокоммуникационные и вычислительные технологии и системы — Улан-Удэ: БГУ, 2010. С. 269 – 272.
9. Эвристический метод обнаружения таблиц в разноформатных документах / И.В. Бычков [и др.] // Вычислительные технологии. 2009. Т. 14, № 2. С. 58-73.
10. Shigarov A.O., Burchkov I.V., Hmelnov A.E., Ruzhnikov G.M. A method for table detection in metafiles // Pattern Recognition and Image Analysis. 2009. Vol. 19, No 4. P. 693-697.
11. Свидетельство об официальной регистрации программы для ЭВМ № 2008614329 «Программная система импорта данных из слабо структурированных таблиц Microsoft Excel в СУБД (Excel2DB)» Бычков И.В., Ружников Г.М., Хмельнов А.Е., Шигаров А.О., 2008.
12. Свидетельство об официальной регистрации программы для ЭВМ № 2008614332 «Программная система извлечения таблиц из электронных документов (TabExtractor)» Бычков И.В., Ружников Г.М., Хмельнов А.Е., Шигаров А.О., 2008.
13. Свидетельство об официальной регистрации программы для ЭВМ № 2008614995 «Рабочие место администратора хранилища статистической информации (AttrTree)» Бычков И.В., Ружников Г.М., Хмельнов А.Е., Шигаров А.О., 2008.
14. Свидетельство об официальной регистрации программы для ЭВМ № 2007613716 «Программное средство извлечения статистических таблиц из неформатированного текста (СтатИмп)» Бычков И.В., Ружников Г.М., Хмельнов А.Е., Шигаров А.О., 2007.

*Ветров Александр Анатольевич*, аспирант ИДСТУ СО РАН, тел. (395-2) 453113, e-mail: [vetrov@icc.ru](mailto:vetrov@icc.ru)

*Фереферов Евгений Сергеевич*, научный сотрудник ИДСТУ СО РАН, тел. (395-2) 453071, e-mail: [fereferov@icc.ru](mailto:fereferov@icc.ru)

*Хмельнов Алексей Евгеньевич*, канд. техн. наук, зав. лаб. ИДСТУ СО РАН, тел. (395-2) 453071, e-mail: [hmelnov@icc.ru](mailto:hmelnov@icc.ru)

*Шигаров Алексей Олегович*, старший научный сотрудник ИДСТУ СО РАН, тел.(395-2) 453102, e-mail: [shigarov@icc.ru](mailto:shigarov@icc.ru)

*Vetrov Alexandr Anatolievich*, PhD student of ISDCT SB RAS.

*Fereferov Evgeniy Sergeevich*, researcher of ISDCT SB RAS.

*Hmelnov Alexey Evgenievich*, candidate of technical science, chief of laboratory of ISDCT SB RAS.

*Shigarov Alexey Olegovich*, doctor of technical science, senior researcher of ISDCT SB RAS.

## МОДЕЛЬ БЕЗОПАСНОСТИ ВЕБ-ПРИЛОЖЕНИЙ НА ОСНОВЕ МАНДАТНОГО РОЛЕВОГО РАЗГРАНИЧЕНИЯ ДОСТУПА<sup>1</sup>

Статья посвящена разработке модели безопасности для веб-приложений. В мандатную ролевую модель добавлены новые элементы, учитывающие функционирование веб-приложений.

**Ключевые слова:** модель безопасности, ролевое разграничение доступа, мандатное разграничение доступа, веб-приложения.

*D.D. Kononov, S.V. Isaev*

### MANDATORY ROLE-BASED ACCESS CONTROL FOR WEB APPLICATIONS

The article is devoted to security model for web-applications. Mandatory role-based access model was extended with new elements related to web applications.

**Keywords:** security model, role-based access control, mandatory access control, web applications.

#### Введение

В связи с развитием систем телекоммуникаций и веб-сервисов вопросы безопасности становятся все более актуальными и востребованными. Целью работы является создание модели безопасности, ориентированной на использование в веб-приложениях.

В настоящее время основными моделями разграничения доступа являются дискреционная, мандатная, ролевая.

Дискреционные модели безопасности основаны на управлении доступа субъектов к объектам с помощью списков управления доступа или матрицы доступа. В этом семействе наиболее исследованы модели Харрисона – Руззо – Ульмана (ХРУ) [1], типизированная матрица доступа [2], Take-Grant [3].

Мандатное разграничение доступа – разграничение доступа субъектов к объектам, основанное на назначении метки конфиденциальности для информации, содержащейся в объектах, и выдаче разрешений субъектам на обращение к информации такого уровня конфиденциальности. Примером мандатной модели является модель Белла – ЛаПадула [4]. В классической модели Белла – ЛаПадулы анализируются условия, при выполнении которых в компьютерной системе невозможно возникновение информационных потоков от объектов с большим уровнем конфиденциальности к объектам с меньшим уровнем конфиденциальности.

Ролевое разграничение доступа является дальнейшим развитием политики дискреционного разграничения доступа: права доступа субъектов к объектам системы группируются по некоторым признакам, образуя роли. Формирование ролей призвано определить четкие и понятные для пользователей правила разграничения доступа.

Перечисленные модели не учитывают специфику функционирования веб-приложений, в частности, иерархическую организацию запросов и ссылок, а также ограничены в способах задания доступа. В работе описана адаптированная модель безопасности, которая устраняет данные проблемы.

#### Модель ролевого разграничения доступа

Базовая модель ролевого разграничения доступа  $RBAC_0$  [5] задается набором элементов:

$\langle U, R, P, S, UA(U), PA(R), user(S), roles(S) \rangle$ ,

где:

U – множество пользователей;

R – множество ролей;

P – множество прав доступа;

S – множество сессий пользователей;

---

<sup>1</sup> Работа выполнена при поддержке интеграционного проекта СО РАН № 21.

UA:  $U \rightarrow 2^R$  – функция, определяющая для каждого пользователя множество ролей, на которые он может быть авторизован;

PA:  $R \rightarrow 2^P$  – функция, определяющая для каждой роли множество прав доступа, при этом  $\forall p \in P, \exists r \in R$  такой, что  $p \in PA(r)$ ;

user:  $S \rightarrow U$  – функция, определяющая для каждой сессии пользователя, от имени которого она авторизована;

roles:  $S \rightarrow 2^R$  – функция, определяющая для пользователя множество ролей, на которые он авторизован в данной сессии; при этом в каждый момент времени  $\forall s \in S$  выполняется условие  $roles(s) \subseteq UA(user(s))$ .

Модель  $RBAC_1$  задается как  $RBAC_0$ , при этом вводится иерархией ролей (RH).

### Адаптация модели для веб-приложений

В модель  $RBAC_1$  к существующим элементам «пользователь» (user), «роль» (role), «разрешение» (permission), «сессия» (session) добавим новые, учитывающие особенности функционирования веб-приложений: «токен» (token), «запрос» (request).

**Определение 1.** Токен ( $Tk$ ) – набор атрибутов пользователя, позволяющих осуществить его аутентификацию в системе. Токеном является пара (имя, пароль) либо пара (сертификат ЭЦП, закрытый ключ ЭЦП).

**Определение 2.** Запрос ( $Rq$ ) – набор информации, пересылаемой клиентом серверу по протоколу HTTP. Запрос содержит набор заголовков, уникальный идентификатор ресурса, набор параметров имя/значение и тело запроса.

Запрос принадлежит сессии, в рамках одной сессии может выполняться несколько запросов. Элементы «запрос» и «разрешение» связаны отношением «многие-ко-многим». На множестве запросов  $Rq$  вводится отношение включения.

**Определение 3.** Запрос  $A$  включает запрос  $B$  ( $B \leq A$ ), если путь уникального идентификатора ресурса (URI) запроса  $A$  содержит путь уникального идентификатора ресурса запроса  $B$  с начальной позиции строки в рамках одного пространства имен, при этом  $len B_{path} \leq len A_{path}$ , где  $len(x)$  – длина строки  $x$ .

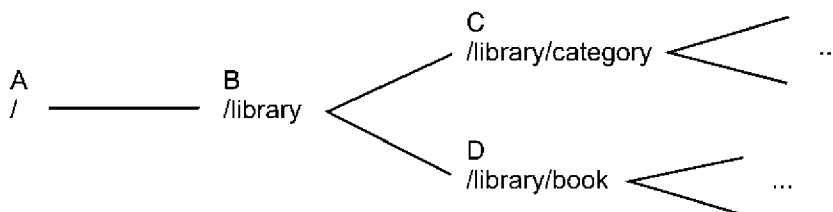


Рис. 1. Пример отношения включения на множестве запросов

На рисунке 1 приведены запросы A, B, C и D, для которых справедливы следующие отношения:  $A \leq B$ ,  $B \leq C$ ,  $B \leq D$ ,  $A \leq C$ ,  $A \leq D$ .

Отношение включения обладает следующими свойствами:

- рефлексивность:

$$\forall rq \in Rq : rq \leq rq,$$

- антисимметричность:

$$\forall rq, rq' \in Rq : ((rq \leq rq') \& (rq' \leq rq)) \rightarrow rq = rq',$$

- транзитивность:

$$\forall rq_1, rq_2, rq_3 \in Rq : ((rq_1 \leq rq_2) \& (rq_2 \leq rq_3)) \rightarrow rq_1 \leq rq_3.$$

Таким образом, отношение включения на множестве запросов  $Rq$  задает отношение нестрогого частичного порядка.

Зададим функцию  $RqA$ , отображающую права доступа на множество запросов  $RqA : P \rightarrow 2^{Rq}$ .

**Определение 4.** Иерархия запросов ( $RqH$ ) – заданное на множестве запросов  $Rq$  отношение включения. При этом для любых  $p \in P$  справедливо условие: если  $rq, rq' \in Rq$ ,  $rq \in RqA(p)$  и  $rq \leq rq'$ , то  $rq' \in RqA(p)$ .

Таким образом, определение 4 дает возможность осуществлять гибкое разграничение доступа для отдельных запросов и всех запросов, которые включают данный запрос.

На рисунке 2 представлена схема элементов полученной модели безопасности.

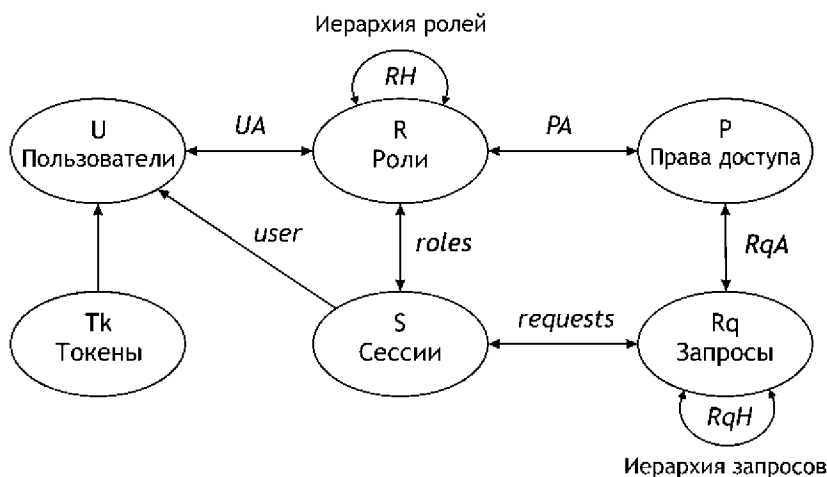


Рис. 2. Схема элементов модели безопасности

#### Адаптированная модель мандатного ролевого разграничения доступа

В работе [6] для базовой ролевой модели описано использование мандатного разграничения доступа, направленной на защиту от угрозы конфиденциальности информации. В рамках введенной терминологии опишем модель мандатного ролевого разграничения доступа для веб-приложений. В дополнение к используемым элементам ролевой модели, заданной выше, добавим следующие элементы:

$Rq$  – множество запросов;

$(L, \leq)$  – решетка уровней конфиденциальности;

$c: U \rightarrow L$  – функция уровней доступа пользователей;

$c: Rq \rightarrow L$  – функция уровней конфиденциальности запросов;

$A = \text{read, write}$  – виды доступа;

$R = \{x\_read \mid x \in L\} \cup \{x\_write \mid x \in L\}$  – множество ролей;

$P = \{(rq, read) \mid rq \in Rq\} \cup \{(rq, write) \mid rq \in Rq\}$  – множество прав доступа.

Таким образом, на множестве ролей  $R$  заданы независимые иерархии по чтению и записи.

Используя определения 5.20 и 5.22 [7], в соответствии с требованиями либерального мандатного управления доступом для множества запросов  $Rq$  задается иерархия на множестве ролей  $R$  и ограничения функций  $UA()$ ,  $roles()$  и  $PA()$ .

В рамках мандатного разграничения доступа задается определение информационного потока.

**Определение 5.** Будем считать, что существует *информационный поток* от запроса  $rq \in Rq$  к запросу  $rq' \in Rq$  тогда и только тогда, когда существуют роли  $r, r' \in R$  и сессия  $s \in S$  такие, что  $(rq, read) \in PA(r)$ ,  $(rq', write) \in PA(r')$  и  $r, r' \in roles(s)$ .

Сформулируем утверждение о невозможности реализации запрещенных информационных потоков от запросов с высоким уровнем конфиденциальности к запросам с низким уровнем конфиденциальности.

**Утверждение 1.** Если модель ролевого управления доступом соответствует требованиям либерального мандатного управления доступом, то в ней для любых запросов  $rq, rq' \in Rq$ , таких что  $c(rq) > c(rq')$ , невозможно возникновение информационного потока от  $rq$  к  $rq'$ .

Доказательство проводится аналогично теореме 5.1 [7].

Таким образом, описанная модель является безопасной с точки зрения информационных потоков для запросов с разным уровнем конфиденциальности.

### Заключение

В работе описана мандатная ролевая модель, учитывающая специфику функционирования веб-приложений. Приведены определения токена, запроса, задано отношение включения на множестве запросов. Сформулировано утверждение о безопасности информационных потоков в системах, построенных на базе данной модели с учетом требований либерального мандатного управления доступом. Проведенное исследование позволило повысить безопасность разрабатываемых веб-приложений [8, 9].

### Литература

1. Harrison M., Ruzzo W. Monotonic protection systems / In DeMillo R., Dobkin D., Jones A., Lipton R., editors // *Foundation of Secure Computation*. New York: Academic Press, 1978. P. 337–365.
2. Sandhu R. The typed access matrix model. In *Proceeding of the IEEE Symposium on Research in Security and Privacy*. Oakland, CA, May 1992. P. 122–136.
3. Теоретические основы компьютерной безопасности / П. Н. Девянин, О. О. Михальский, Д. И. Правиков, А. Ю. Щербаков. М.: Радио и связь, 2000.
4. Bell D. E., LaPadula L. J. *Secure Computer Systems: Unified Exposition and Multics Interpretation*. – Bedford, Mass.: MITRE Corp., 1976. – MTR-2997 Rev. 1.
5. Sandhu R., Coyne E. J., Feinstein H. L. and Youman C. E. Role-based Access Control Models // *IEEE Computer* (IEEE Press), vol. 29. No. 2, 1996.
6. Sandhu R. Role-based Access Control // *Advanced computers*, vol. 46, 1998. P. 237–286.
7. Девянин П. Н. Модели безопасности компьютерных систем: учеб. пособие для студ. высш. учеб. заведений. М.: Академия, 2005. – 144 с.
8. Кононов Д. Д., Исаев С. В., Исаева О. С. Организация информационных интернет-ресурсов для поддержки управления муниципальным заказом // *Вестник СибГАУ*, 2009. Вып. 1 (22), ч. 2. С. 35–40.
9. Кононов Д. Д., Исаев С. В. Модель безопасности кросс-платформенных веб-сервисов поддержки муниципальных закупок // *Прикладная дискретная математика*. Приложение. 2011. № 4. С. 48–50.

*Кононов Дмитрий Дмитриевич*, научный сотрудник Института вычислительного моделирования СО РАН, 660036, г. Красноярск, Академгородок, 50/44, тел. (391) 2494767, e-mail: ddk@icm.krasn.ru

*Исаев Сергей Владиславович*, канд. техн. наук, старший научный сотрудник Института вычислительного моделирования СО РАН, 660036, г. Красноярск, Академгородок, 50/44, тел. (391) 2494767, e-mail: i@icm.krasn.ru

*Kononov Dmitry Dmitrievich*, scientific researcher of Institute of Computational Modeling SB RAS.

*Isaev Sergey Vladislavovich*, candidate of technical sciences, senior scientific researcher of Institute of Computational Modeling SB RAS.

## МОДЕЛИРОВАНИЕ ПОИСКА РЕШЕНИЯ С ПОМОЩЬЮ ФУНКЦИОНАЛЬНЫХ ГРАММАТИК

Статья посвящена проблеме автоматизированного поиска решения задач. Рассматривается моделирование решения в области знаний, представленной в виде функциональной грамматики, а также возможность реализации модели на функциональном языке программирования ЛИСП. Показываются сильные стороны использования аппарата функциональных грамматик.

**Ключевые слова:** грамматика, ЛИСП, поиск решения, синтез программ.

V.A. Kravchenko

## MODELLING OF SEARCH OF THE DECISION BY FUNCTIONAL GRAMMARS

Article is devoted a problem of computer-assisted retrieval of the decision of tasks. Decision modeling in the knowledge domain presented in the form of the functional grammar, and also possibility of implementation of model in the functional programming language LISP is considered. Advantages of usage of functional grammars are shown.

**Keywords:** grammar, LISP, decision search, synthesis of programs.

### Введение

Целью работы является моделирование поиска решения задач в конкретной предметной области и его программная реализация. Система решения должна состоять из базы знаний и механизма вывода (рис. 1). База знаний предназначена для хранения правил теории, в рамках которой требуется решение задач. Чем больше правил содержится в базе знаний, тем больший круг задач и с большей эффективностью она способна решать. Задачей механизма вывода является способность из фактов, представляющих собой исходную информацию, посредством правил, содержащихся в базе знаний, получить в качестве результата нужную информацию. Следует отметить, что система должна иметь обучающийся характер. Для этого систему необходимо снабдить способностью преобразовывать результат в правила системы, пополняя тем самым её базу знаний.

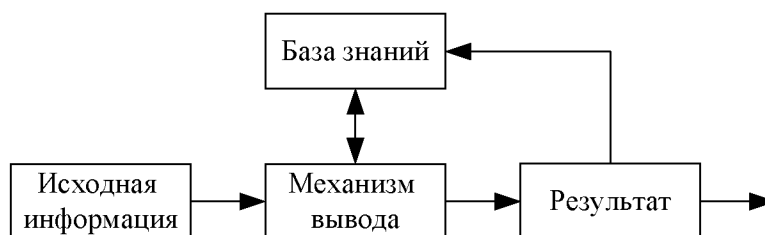


Рис. 1. Система поиска решения

В данной работе предлагается использовать функциональный подход. Он заключается в моделировании системы аппаратом функциональных грамматик и её реализацию с помощью функционального языка программирования ЛИСП.

### 1. Функциональные грамматики

В качестве функциональной грамматики будем понимать формальную грамматику вида:

$$G = (V_N, V_T, F, P, S),$$

где  $V_N$  – конечный алфавит нетерминальных символов;  $V_T$  – конечный алфавит терминальных символов;  $F$  – конечное множество функций  $f$ ;  $P$  – конечное множество правил вида:  $A \rightarrow \varphi \{f\}$ , где  $A \in V_N$ ,  $\varphi \in (V_N \cup V_T)^*$ ,  $f \in F$ ;  $S$  – начальный нетерминальный символ (аксиома),  $S \in V_N$ .

Такая грамматика может быть получена из любой формальной грамматики Хомского второго типа:

$$G1 = (V_N, V_{T1}, P1, S),$$

где  $V_N$  – конечный алфавит нетерминальных символов;  $V_{T1}$  – конечный алфавит терминальных символов;  $P$  – конечное множество правил вида:  $A \rightarrow \psi$ , где  $A \in V_N$ ,  $\psi \in (V_N \cup V_{T1})^*$ ,  $S$  – начальный нетерминальный символ,  $S \in V_N$ .

Для перехода от КС-грамматики к функциональной грамматике необходимо из всего алфавита терминальных символов  $V_{T1}$  выделить терминальные символы  $V_{T0} \in V_{T1}$ , определяемые через базисные функции  $f0$  рассматриваемой системы. Тогда алфавит терминальных символов сократиться до множества  $V_T$ , причем  $V_{T1} \equiv (V_T \cup V_{T0})$ . В результате сокращения алфавита символов будут упрощены правые части продукций  $A \rightarrow \psi$ . Для сохранения синтаксиса и семантики языка грамматики необходимо каждой упрощенной продукции  $A \rightarrow \phi$  поставить в соответствие функцию  $f$ , аргументами которой будут являться все символы последовательности  $\phi$ . При этом функция  $f$  должна быть представлена как суперпозиция базисных функций  $f0$ . Тогда множество правил грамматики  $P$  будет иметь вид:  $A \rightarrow \phi \{f\}$ , где  $A \in V_N$ ,  $\phi \in (V_N \cup V_T)^*$ ,  $f \in F$ , причём также  $f0 \in F0$ ,  $F0 \subseteq F$ .

Рассмотрим процесс перехода к функциональной грамматике на небольшом примере. Пусть имеется КС-грамматика:

$$G1 = (V_N, V_{T1}, P1, S);$$

$$V_N = S, A, B, C, D, E ; \quad V_{T1} = a, b, c, d, e, \alpha, \beta, \gamma ; \quad S = S;$$

$$P1 = \left\{ \begin{array}{l} S \rightarrow A\alpha\beta B; \\ A \rightarrow b\beta C\alpha c \mid \gamma d\beta e\alpha D; \\ B \rightarrow \gamma a\beta b\alpha c \mid d\alpha e\beta a\delta; \\ C \rightarrow \gamma E\alpha b; \\ D \rightarrow c\beta E\delta; \\ E \rightarrow \gamma d\alpha e. \end{array} \right\}.$$

Преобразуем заданную грамматику к функциональному виду. Для этого выберем в качестве базисных символов  $V_{T0} = \alpha, \beta, \gamma$ . Тогда схема функциональной грамматики будет иметь вид:

$$G = (V_N, V_T, F, P, S);$$

$$V_N = S, A, B, C, D, E ; \quad V_T = a, b, c, d, e ; \quad S = S;$$

$$P = \left\{ \begin{array}{l} S \rightarrow AaB \{f_1\}; \\ A \rightarrow bCc \{f_2\} \mid deD \{f_3\}; \\ B \rightarrow abc \{f_3\} \mid dea \{f_4\}; \\ C \rightarrow Eb \{f_5\}; \\ D \rightarrow cE \{f_6\}; \\ E \rightarrow de \{f_3\}; \end{array} \right\},$$

$$F = \left\{ \begin{array}{ll} f_1 = (x, y, z) : x\alpha y\beta z; & f0_1 = (x, y) : x\alpha y; \\ f_2 = (x, y, z) : x\beta y\alpha z; & f0_2 = (x, y) : x\beta y; \\ f_3 = (x, y, z) : \gamma x\beta y\alpha z; & f0_3 = (x) : \gamma x; \\ f_4 = (x, y, z) : x\alpha y\beta z\delta; & f0_4 = (x) : x\delta; \\ f_5 = (x, y) : \gamma x\alpha y; & \\ f_6 = (x, y) : x\beta y\delta; & \end{array} \right\}.$$

Обратим внимание, что функции  $f0_1..f0_4$  образуют подсистему базисных функций, а функции  $f_1..f_6$  выражаются через суперпозиции этих базисных функций:



$$\begin{aligned}
 f_1(x, y, z) &= f_{0_2}(f_{0_1}(x, y), z); \\
 f_2(x, y, z) &= f_{0_1}(f_{0_2}(x, y), z); \\
 f_3(x, y, z) &= f_{0_1} f_{0_2} f_{0_3} x, y, z; \\
 f_4(x, y, z) &= f_{0_4} f_{0_2} f_{0_1} x, y, z; \\
 f_5(x, y) &= f_{0_1} f_{0_3} x, y; \\
 f_6(x, y) &= f_{0_4} f_{0_2} x, y.
 \end{aligned}$$

Построим для сравнения два дерева синтаксического разбора для грамматики G1 (рис. 2) и грамматики G (рис. 3). Ветви второго дерева содержат функции, соответствующие применяемым продукциям. В результате этого, каждому выводу дерева функциональной грамматики соответствует суперпозиция функций:

$$\begin{aligned}
 \sigma_1 &= f_1(f_2(b, f_5(f_5(d, e), b)), c), a, f_3(a, b, c)); \\
 \sigma_2 &= f_1(f_2(b, f_5(f_5(d, e), b)), c), a, f_4(d, e, a)); \\
 \sigma_3 &= f_1(f_3(d, e, f_6(c, f_5(d, e))), a, f_3(a, b, c)); \\
 \sigma_4 &= f_1(f_3(d, e, f_6(c, f_5(d, e))), a, f_4(d, e, a)).
 \end{aligned}$$

На основе сравнения схем и деревьев грамматик, можно сделать вывод, что функциональные грамматики имеют ряд особенностей по сравнению с традиционными КС-грамматиками.

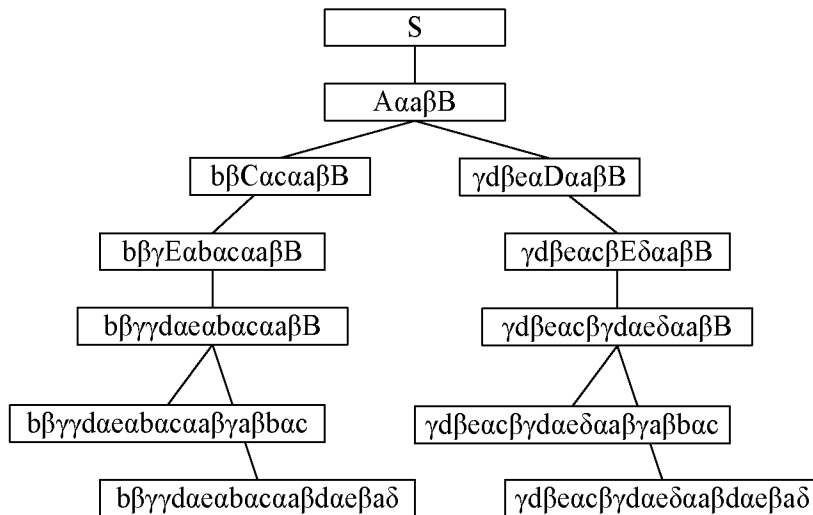


Рис. 2. Дерево разбора для грамматики G1

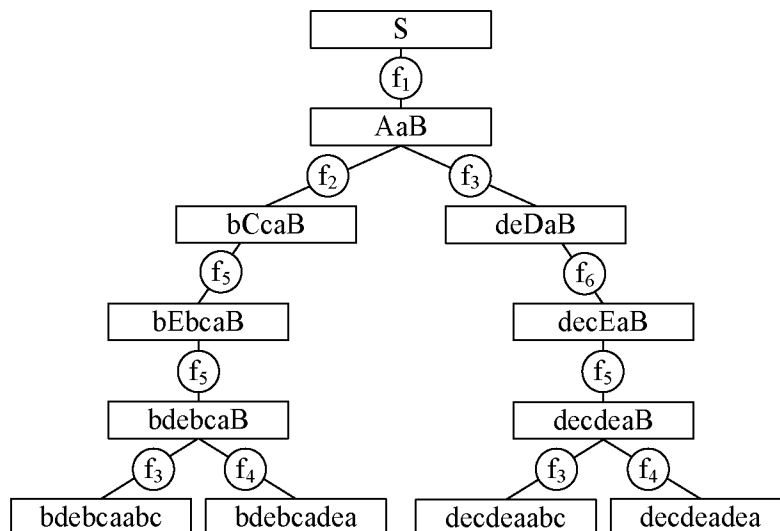


Рис. 3. Дерево разбора для грамматики G

Во-первых, схема функциональной грамматики имеет более структурированный вид. Во-

вторых, цепочки порожденного языка являются более короткими ввиду сокращенного алфавита терминальных символов. В-третьих, каждому выводу в функциональной грамматике соответствует суперпозиция функций, позволяющая проследить его логику.

## 2. База знаний

База знаний системы поиска решения должна содержать правила, описывающие все доступные законы, зависимости и факты рассматриваемой предметной области. При этом весьма актуальным является проблема представления знаний. Рассмотрим возможность представления знаний с помощью описанных выше функциональных грамматик.

Для того, чтобы создать базу знаний, необходимо задать объединенный алфавит грамматики  $V \equiv (V_N \cup V_T)^*$ , систему функций  $F$  и систему выраженных посредством них правил  $P$ . В качестве символов грамматики используются метапеременные, обозначающие все объекты рассматриваемой предметной области. С помощью правил системы  $P$  в виде:  $\alpha \rightarrow \varphi \{f\}$  где  $A \in V$ ,  $\varphi \in V$ ; и функций  $f \in F$  должны быть записаны все известные закономерности рассматриваемой теории. Важно напомнить, что система функций  $F$  обязательно должна содержать подсистему  $F_0$  базисных функций.

Рассмотрим в качестве примера базу знаний по механике прямолинейного равноускоренного движения тела. В качестве алфавита грамматики выбираем 14 величин: <Dist> – расстояние; <Time> – время; <Spd1> – начальная скорость; <Spd2> – конечная скорость; <Accl> – ускорение; <Mass> – масса; <KinEn1> – начальная кинетическая энергия; <KinEn2> – конечная кинетическая энергия; <FrPull> – сила тяги; <FrFric> – сила трения; <OpPull> – работа силы тяги; <OpFric> – работа силы трения; <CfFric> – коэффициент трения; <Cfg> – ускорение свободного падения. В качестве алфавита символов, выражающих базисные функции, определяем <+>, <->, <\*>, </>, <^2>, <КОРЕНЬ>.

С учетом основных закономерностей прямолинейного равноускоренного движения, можно составить системы продукций и соответствующих им функций. Грамматика примет вид:

$$G = (V_N, V_T, F, P, S);$$

$$V = \left\{ \begin{array}{l} \langle \text{Dist} \rangle, \langle \text{Time} \rangle, \langle \text{Spd1} \rangle, \langle \text{Spd2} \rangle, \langle \text{Accl} \rangle, \\ \langle \text{Mass} \rangle, \langle \text{KinEn1} \rangle, \langle \text{KinEn2} \rangle, \langle \text{FrPull} \rangle, \\ \langle \text{FrFric} \rangle, \langle \text{OpPull} \rangle, \langle \text{OpFric} \rangle, \langle \text{CfFric} \rangle, \langle \text{Cfg} \rangle \end{array} \right\};$$

$$P = \{ \begin{array}{l} \langle \text{Spd2} \rangle \rightarrow \langle \text{Spd1} \rangle \langle \text{Accl} \rangle \langle \text{Time} \rangle \{f_1\}; \\ \langle \text{Spd2} \rangle \rightarrow \langle \text{Accl} \rangle \langle \text{Dist} \rangle \langle \text{Spd1} \rangle \{f_2\}; \\ \langle \text{Spd2} \rangle \rightarrow \langle \text{KinEn2} \rangle \langle \text{Mass} \rangle \{f_{11}\}; \\ \langle \text{Spd1} \rangle \rightarrow \langle \text{Spd2} \rangle \langle \text{Accl} \rangle \langle \text{Time} \rangle \{f_3\}; \\ \langle \text{Spd1} \rangle \rightarrow \langle \text{Dist} \rangle \langle \text{Time} \rangle \langle \text{Accl} \rangle \{f_4\}; \\ \langle \text{Spd1} \rangle \rightarrow \langle \text{Spd2} \rangle \langle \text{Accl} \rangle \langle \text{Dist} \rangle \{f_5\}; \\ \langle \text{Spd1} \rangle \rightarrow \langle \text{KinEn1} \rangle \langle \text{Mass} \rangle \{f_{11}\}; \\ \langle \text{Dist} \rangle \rightarrow \langle \text{Spd1} \rangle \langle \text{Time} \rangle \langle \text{Accl} \rangle \langle \text{Time} \rangle \{f_6\}; \\ \langle \text{Dist} \rangle \rightarrow \langle \text{Spd2} \rangle \langle \text{Spd1} \rangle \langle \text{Accl} \rangle \{f_7\}; \\ \langle \text{Dist} \rangle \rightarrow \langle \text{OpPull} \rangle \langle \text{FrPull} \rangle \{f_{12}\}; \\ \langle \text{Dist} \rangle \rightarrow \langle \text{OpFric} \rangle \langle \text{FrFric} \rangle \{f_{12}\}; \\ \langle \text{Time} \rangle \rightarrow \langle \text{Spd2} \rangle \langle \text{Spd1} \rangle \langle \text{Accl} \rangle \{f_8\}; \\ \langle \text{Time} \rangle \rightarrow \langle \text{Spd1} \rangle \langle \text{Accl} \rangle \langle \text{Dist} \rangle \{f_9\}; \\ \langle \text{Accl} \rangle \rightarrow \langle \text{Spd2} \rangle \langle \text{Spd1} \rangle \langle \text{Time} \rangle \{f_8\}; \\ \langle \text{Accl} \rangle \rightarrow \langle \text{Dist} \rangle \langle \text{Spd1} \rangle \langle \text{Time} \rangle \{f_{10}\}; \\ \langle \text{Accl} \rangle \rightarrow \langle \text{Spd2} \rangle \langle \text{Spd1} \rangle \langle \text{Dist} \rangle \{f_7\}. \\ \langle \text{Accl} \rangle \rightarrow \langle \text{FrPull} \rangle \langle \text{FrFric} \rangle \langle \text{Mass} \rangle \{f_8\}; \\ \langle \text{Mass} \rangle \rightarrow \langle \text{FrPull} \rangle \langle \text{FrFric} \rangle \langle \text{Accl} \rangle \{f_8\}; \\ \langle \text{Mass} \rangle \rightarrow \langle \text{FrFric} \rangle \langle \text{CfFric} \rangle \langle \text{Cfg} \rangle \{f_{13}\}; \\ \langle \text{Mass} \rangle \rightarrow \langle \text{KinEn2} \rangle \langle \text{Spd2} \rangle \{f_{14}\}; \\ \langle \text{Mass} \rangle \rightarrow \langle \text{KinEn1} \rangle \langle \text{Spd1} \rangle \{f_{14}\}; \\ \langle \text{FrPull} \rangle \rightarrow \langle \text{FrFric} \rangle \langle \text{Mass} \rangle \langle \text{Accl} \rangle \{f_1\}; \\ \langle \text{FrPull} \rangle \rightarrow \langle \text{OpPull} \rangle \langle \text{Dist} \rangle \{f_{12}\}; \\ \langle \text{FrFric} \rangle \rightarrow \langle \text{FrPull} \rangle \langle \text{Mass} \rangle \langle \text{Accl} \rangle \{f_3\}; \end{array} \}$$

$\langle \text{FrFric} \rangle \rightarrow \langle \text{Cfg} \rangle \langle \text{CfFric} \rangle \langle \text{Mass} \rangle \{f_{15}\};$   
 $\langle \text{FrFric} \rangle \rightarrow \langle \text{OpFric} \rangle \langle \text{Dist} \rangle \{f_{12}\};$   
 $\langle \text{CfFric} \rangle \rightarrow \langle \text{FrFric} \rangle \langle \text{Mass} \rangle \langle \text{Cfg} \rangle \{f_{13}\};$   
 $\langle \text{KinEn2} \rangle \rightarrow \langle \text{Mass} \rangle \langle \text{Spd2} \rangle \{f_{16}\};$   
 $\langle \text{KinEn2} \rangle \rightarrow \langle \text{OpPull} \rangle \langle \text{OpFric} \rangle \langle \text{KinEn1} \rangle \{f_{17}\};$   
 $\langle \text{KinEn1} \rangle \rightarrow \langle \text{Mass} \rangle \langle \text{Spd1} \rangle \{f_{16}\};$   
 $\langle \text{KinEn1} \rangle \rightarrow \langle \text{KinEn2} \rangle \langle \text{OpPull} \rangle \langle \text{OpFric} \rangle \{f_{17}\};$   
 $\langle \text{OpPull} \rangle \rightarrow \langle \text{FrPull} \rangle \langle \text{Dist} \rangle \{f_{18}\};$   
 $\langle \text{OpPull} \rangle \rightarrow \langle \text{KinEn2} \rangle \langle \text{KinEn1} \rangle \langle \text{OpFric} \rangle \{f_{17}\};$   
 $\langle \text{OpFric} \rangle \rightarrow \langle \text{FrFric} \rangle \langle \text{Dist} \rangle \{f_{18}\};$   
 $\langle \text{OpFric} \rangle \rightarrow \langle \text{OpPull} \rangle \langle \text{KinEn2} \rangle \langle \text{KinEn1} \rangle \{f_{17}\} \quad \}.$

$F = \{$   
 $f_0 = (x, y): x + y;$   
 $f_1 = (x, y): x - y;$   
 $f_2 = (x, y): x * y;$   
 $f_3 = (x, y): x / y;$   
 $f_4 = (x): x^2;$   
 $f_5 = (x): \text{КОРЕНЬ}(x);$   
 $f_6 = (x, y, z): f_0(x, f_3(y, z));$   
 $f_7 = (x, y, z): f_6(f_0(f_3(f_3(x, y), 2), f_5(z)));$   
 $f_8 = (x, y, z): f_2(x, f_3(y, z));$   
 $f_9 = (x, y, z): f_2(f_4(x, y), f_4(f_3(f_5(y), z) 2));$   
 $f_{10} = (x, y, z): f_6(f_2(f_5(x), f_3(f_3(y, z), 2)));$   
 $f_{11} = (x, y, z): f_0(f_3(x, y), f_4(f_3(f_5(y), z), 2));$   
 $f_{12} = (x, y, z): f_4(f_2(f_5(x), f_5(y)), f_3(2, z));$   
 $f_{13} = (x, y, z): f_4(f_2(x, y), z);$   
 $f_{14} = (x, y, z): f_4(f_2(f_0(f_5(x), f_5(f_3(y, z), 2))), x, y);$   
 $f_{15} = (x, y, z): f_4(f_3(2, f_2(x, f_3(y, z))), f_5(z));$   
 $f_{16} = (x, y, z): f_6(f_4(f_3(2, x), y));$   
 $f_{17} = (x, y): f_4(x, y);$   
 $f_{18} = (x, y, z): f_4(f_3(x, y), z);$   
 $f_{19} = (x, y): f_4(f_3(2, x), f_5(y));$   
 $f_{20} = (x, y, z): f_3(f_3(x, y), z);$   
 $f_{21} = (x, y): f_4(f_3(x, f_5(y)), 2);$   
 $f_{22} = (x, y, z): f_0(f_2(x, y), z);$   
 $f_{23} = (x, y): f_3(x, y) \quad \}.$

Таким образом, была составлена база знаний, хотя записанную грамматику нельзя считать полной, так как в ней не разграничены терминальные и нетерминальные символы, не задан начальный символ. Эти элементы определяются при решении конкретной задачи на стадии получения исходных данных. Базу знаний в любой момент можно дополнить, внося в грамматику новые правила и функции.

### 3. Механизм вывода

Механизм вывода предназначен для нахождения результата по исходным данным посредством правил, содержащихся в базе знаний.

При решении конкретной задачи, перед поиском решения необходимо окончательно сформировать функциональную грамматику, разграничив терминальные и нетерминальные символы и задав аксиому. В качестве последней выбирается величина, которую нужно найти, в качестве терминалов – известные из условия величины, а в качестве нетерминалов – все остальные символы, входящие в общий алфавит. Нетерминальные символы, тем самым, отражают неизвестные величины, нахождение которых не требуется.

Таким образом, для решения каждой задачи индивидуально задаются аксиома, терминальные и нетерминальные символы, т.е. грамматика приобретает завершённый вид. Продукции и общий алфавит постоянны в рамках предметной области. Кроме того, следует заметить, что если неизвестных величин в задаче несколько, то решение разбивается на несколько этапов, в каждом из которых заново задается аксиома. Аксиомы предыдущих этапов следует в качестве известных величин переводить в класс терминальных символов.

В рамках определенной функциональной грамматики появляется возможность построить дерево разбора. Оно необходимо для поиска полного вывода, соединяющего корневой узел – аксиому – с узлом, представленным только терминальными символами. Для этого последовательно применяются продукции грамматики, которые позволяют заменять нетерминалы на цепочки левых частей продукций. Если в результате такой замены появляются аксиома или же замененные ранее нетерминалы, то такие узлы считаются тупиковыми и должны быть отброшены. В результате применения каждой продукции образуется новая ветвь дерева, которой сопоставляется функция данной продукции. Поэтому найденному полному выводу будет соответствовать суперпозиция функций, расположенных в узлах, через которые проходит этот вывод. Данная суперпозиция и отражает решение задачи.

Стоит отметить, что в одном дереве разбора вполне может оказаться несколько полных выводов, поэтому существует задача выбора более простого решения, содержащего минимальное количество функций. Для этого должен применяться поиск «в ширину». Такой способ подразумевает изучение всех состояний каждого уровня, и только после этого – переход на более глубокий уровень. Каждый новый уровень означает применение новой функции, поэтому более глубокие решения являются более сложными.

Рассмотрим поиск решения на примере заданной выше базы знаний. Пусть в задаче известны: сила тяги, действующая на тело, пройденный им путь, его начальная и конечная кинетические энергии, а требуется найти действующую силу трения. В этом случае, роль терминальных символов будут играть: <FrPull>, <Dist>, <KinEn1>, <KinEn2>; а аксиомы – <FrFric>. Дерево разбора для такой задачи при использовании поиска «в ширину» имеет вид, представленный на рисунке 4.

Решение было найдено на третьем уровне дерева поиска. Ему соответствует следующая суперпозиция функций:

$$\langle \text{FrFric} \rangle = f_{12}(f_{17}(f_{18}(\langle \text{FrPull} \rangle, \langle \text{Dist} \rangle), \langle \text{KinEn2} \rangle, \langle \text{KinEn1} \rangle), \langle \text{Dist} \rangle).$$

#### 4. Реализация на языке ЛИСП

Язык ЛИСП идеально подходит для реализации базы знаний и механизма вывода, представленных аппаратом функциональных грамматик, так как обладает возможностью символьной обработки и является функциональным языком программирования. Программа на ЛИСПе представляет собой суперпозицию функций, и результат рассмотренной системы решений также является суперпозицией функций, поэтому при грамотной реализации полученный результат будет являться программой. Таким образом, ЛИСП позволяет не только осуществлять поиск решения, но производить синтез программ.

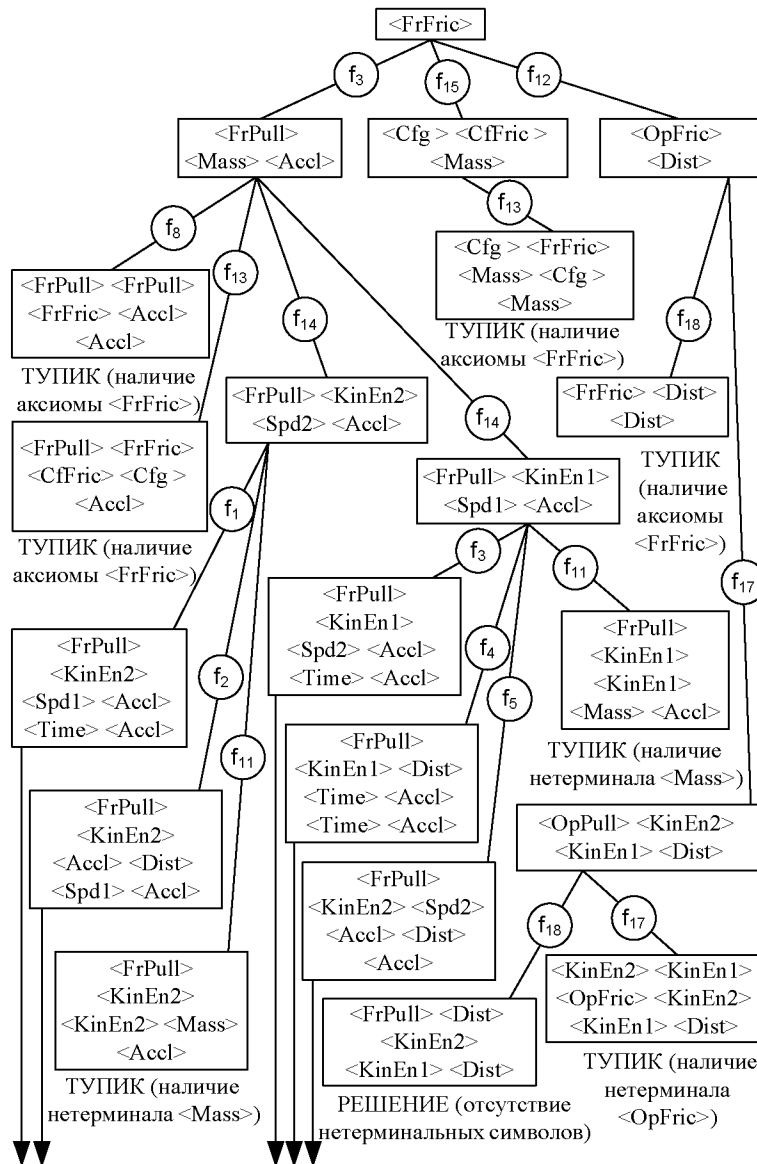


Рис. 4. Дерево разбора поиска решения задачи

ЛИСП при своей работе оперирует списками, представляющими собой упорядоченные совокупности атомов и подписков. База знаний на ЛИСПе может быть представлена списком ПРОДУКЦИИ и функциями  $f_1..f_n$ . Для рассмотренной выше базы знаний список ПРОДУКЦИИ имеет вид:

- ((<OpFric> (f17 <OpPull> <KinEn2> <KinEn1>) (f18 <FrFric> <Dist>))
- (<OpPull> (f17 <KinEn2> <KinEn1> <OpFric>) (f18 <FrPull> <Dist>))
- (<KinEn1> (f16 <Mass> <Spd1>) (f17 <KinEn2> <OpPull> <OpFric>))
- (<KinEn2> (f16 <Mass> <Spd2>) (f17 <OpPull> <OpFric> <KinEn1>))
- (<CfFric> (f13 <FrFric> <Mass>))
- (<FrFric> (f12 <OpFric> <Dist>) (f15 <Cfg> <CfFric> <Mass>) (f3 <FrPull> <Mass> <Accl>))
- (<FrPull> (f12 <OpPull> <Dist>) (f1 <FrFric> <Mass> <Accl>))
- (<Mass> (f14 <KinEn1> <Spd1>) (f14 <KinEn2> <Spd2>) (f13 <FrFric> <CfFric> <Cfg>) (f8 <FrPull> <FrFric> <Accl>))
- (<Accl> (f8 <FrPull> <FrFric> <Mass>) (f7 <Spd2> <Spd1> <Dist>) (f10 <Dist> <Spd1> <Time>) (f8 <Spd2> <Spd1> <Time>))
- (<Time> (f9 <Spd1> <Accl> <Dist>) (f8 <Spd2> <Spd1> <Accl>))
- (<Dist> (f12 <OpFric> <FrFric>) (f12 <OpPull> <FrPull>) (f7 <Spd2> <Spd1> <Accl>) (f6 <Spd1> <Time> <Accl>))
- (<Spd1> (f11 <KinEn1> <Mass>) (f5 <Spd2> <Accl> <Dist>) (f4 <Dist> <Time> <Accl>) (f3 <Spd2> <Accl> <Time>))

(<Spd2> (f11 <KinEn2> <Mass>) (f2 <Accl> <Dist> <Spd1>) (f1 <Spd1> <Accl> <Time>)))

Функции грамматики рассматриваемой базы знаний:

```
(defun f1 (x y z) (+ x (* y z)))
(defun f2 (x y z) (sqr (+ (* 2 x y) (* z z))))
(defun f3 (x y z) (- x (* y z)))
(defun f4 (x y z) (- (/ x y) (/ (* y y z) 2)))
(defun f5 (x y z) (sqr (- (* x x) (* 2 y z))))
(defun f6 (x y z) (+ (* x y) (/ (* y y z) 2)))
(defun f7 (x y z) (/ (- (* x x) (* y y)) (* 2 z)))
(defun f8 (x y z) (/ (- x y) z))
(defun f9 (x y z) (/ (- (sqr (+ (* x x) (* 2 y z) x)) y) y))
(defun f10 (x y z) (/ (* 2 (- x (* y z))) (* z z)))
(defun f11 (x y) (sqr (/ (* 2 x) y)))
(defun f12 (x y) (/ x y))
(defun f13 (x y z) (/ (* x y) z))
(defun f14 (x y) (/ (* 2 x) (* y y)))
(defun f15 (x y z) (* x y z))
(defun f16 (x y) (/ (* x y y) 2))
(defun f17 (x y z) (+ (- x y) z))
(defun f18 (x y) (* x y))
```

При поступлении исходных данных формируется списки ПРОДУКЦИИ1, СИМВОЛЫ и ДЕРЕВО.

Список СИМВОЛЫ состоит из терминальных символов:

(<FrPull> <Dist> <KinEn1> <KinEn2>)

Список ДЕРЕВО состоит из одного атома – начального нетерминального символа:

(<FrFric>)

Список ПРОДУКЦИИ1 представляет собой сокращенный список ПРОДУКЦИИ, избавленный от «неправильных» правил, содержащих в левой части терминальные символы. Его построение осуществляется при участии списка СИМВОЛЫ.

Список ДЕРЕВО является динамическим. Сначала он разрастается за счет применения продукций, а потом, в случае наличия решения, уменьшается за счёт отсекаания тупиковых ветвей. Решением задачи будет являться список ДЕРЕВО, представленный суперпозицией функций fi.

Рассмотрим порядок работы ЛИСПа со списком ДЕРЕВО в ходе вывода решения. Поиск решения состоит из двух этапов, повторяющихся на каждом новом уровне дерева. Первый этап – вставка продукций, второй – анализ новых узлов.

Первый этап заключается в поиске подсписков, состоящих только из атомов и замене самых левых нетерминалов таких подсписков на соответствующие правила, взятые из списка ПРОДУКЦИИ1. При этом определение нетерминалов осуществляется с помощью списка СИМВОЛЫ.

Второй этап включает в себя 2 действия:

1. поиск подсписка состоящего только из терминалов и, в случае успеха, выделение этого подсписка в качестве окончательного списка ДЕРЕВО и прекращение процесса решения;
2. поиск подсписков, содержащих аксиому или ранее замененный нетерминальный символ, удаление таких подсписков и перевод процесса на первый этап, т.е. на следующий уровень дерева.

При поиске решения, представленном на рисунке 4, список ДЕРЕВО будет изменяться в следующем порядке:

Первоначальное состояние:

(<FrFric>)

Уровень 1, этап 1 – вставка продукций <FrFric>:

(<FrFric> (f12 <OpFric> <Dist>) (f15 <Cfg> <CfFric> <Mass>) (f3 <FrPull> <Mass> <Accl>))

Уровень 1, этап 2 – без изменений:

(<FrFric> (f12 <OpFric> <Dist>) (f15 <Cfg> <CfFric> <Mass>) (f3 <FrPull> <Mass> <Accl>))

Уровень 2, этап 1 – вставка продукций <Mass>, <CfFric>, <OpFric>:

(<FrFric> (f12 (<OpFric> (f17 <OpPull> <KinEn2> <KinEn1>) (f18 <FrFric> <Dist>)) <Dist>) (f15 <Cfg> (<CfFric> (f13 <FrFric> <Mass>)) <Mass>) (f3 <FrPull> (<Mass> (f14 <KinEn1> <Spd1>)(f14 <KinEn2> <Spd2>)) (f13 <FrFric> <CfFric> <Cfg>) (f8 <FrPull> <FrFric> <Accl>)) <Accl>))

Уровень 2, этап 2 – удаление 4 тупиковых ветвей:

```
(<FrFric> (f12 (<OpFric> (f17 <OpPull> <KinEn2> <KinEn1>)) <Dist>) (f3 <FrPull> (<Mass> (f14 <KinEn1> <Spd1>) (f14 <KinEn2> <Spd2>))<Accl>))
```

Уровень 3, этап 1 – вставка продукций <Spd2>, <Spd1>, <OpPull>:

```
(<FrFric> (f12 (<OpFric> (f17 (<OpPull> (f17 <KinEn2> <KinEn1> <OpFric>) (f18 <FrPull> <Dist>)) <KinEn2> <KinEn1>)) <Dist>) (f3 <FrPull> (<Mass> (f14 <KinEn1> (<Spd1> (f11 <KinEn1> <Mass>)) (f5 <Spd2> <Accl> <Dist>)) (f4 <Dist> <Time> <Accl>)) (f3 <Spd2> <Accl> <Time>))) (f14 <KinEn2> (<Spd1> (f11 <KinEn1> <Mass>)) (f5 <Spd2> <Accl> <Dist>)) (f4 <Dist> <Time> <Accl>)) (f3 <Spd2> <Accl> <Time>)))<Accl>))
```

Уровень 3, этап 2 – выделение решения:

```
(f12 (f17 (f18 <FrPull> <Dist>) <KinEn2> <KinEn1>) <Dist>)
```

Таким образом, результатом является всё та же суперпозиция функций, но записанная на ЛИСПе в качестве списка ДЕРЕВО. Существует несколько вариантов использования полученного результата:

1. вывести суперпозицию функций в понятном для пользователя виде для демонстрации хода решения поставленной им задачи;

2. выполнить суперпозицию функций для получения числового результата, предварительно задав числовые значения исходных данных;

3. использовать суперпозицию функций для синтеза программы на каком-либо языке программирования;

4. создать из суперпозиции функций новую продукцию, пополнив базу знаний.

Выполнение первых трех действий зависит от цели поиска решения, а выполнение четвертого действия необходимо при каждом решении для накопления системой опыта.

### Заключение

Использование в системе решений функциональных грамматик позволяет смоделировать базу знаний в виде множества функций, а процесс решения и его результат – в виде упорядоченной комбинации таких функций. Это приводит к получению ряда преимуществ.

Во-первых, база знаний может оперативно пополняться за счет добавление новых функций. Причем пополнение знаний возможно в результате накопления опыта, т.к. результат решения также представляет собой функцию. Новые продукции могут быть выражены через уже имеющиеся функции, что упрощает процесс добавления знаний и позволяет избежать повторений, пересечений и противоречий в правилах.

Во-вторых, дерево разбора функциональной грамматики идеально подходит для моделирования поиска решения. Оперирование понятиями функциональной грамматики можно точно описать принцип поиска, что необходимо для его программной реализации.

В-третьих, представление результата в виде суперпозиции функций позволяет применять его для разных задач, в том числе для синтеза программы. Суперпозиция позволяет ответить не только на вопрос, каков результат, но и на вопрос, как этот результат достичь. Действительно, по найденной суперпозиции всегда можно восстановить ветвь решения от исходных данных до результата.

Реализация системы решений на функциональном языке символьного программирования ЛИСП дает возможность полностью использовать все возможности функциональных грамматик. Это позволяет избежать лишних надстроек и пополнять базу знаний без какого-либо вмешательства в механизм вывода.

### Литература

1. Тузов В.А. Математическая модель языка. – Л.: Изд-во Ленингр. ун-та, 1984. – 176 с.
2. Кравченко В. А., Могнонов П. Б., Чимитов Д. Н. Представление знаний в функциональных грамматиках // Вестник СибГАУ. – Красноярск: Изд-во СибГАУ, 2011. №5 (38). С. 55-61.
3. Хьювёнен Э., Сеппянен И. Мир Лиспа. Т.2: Методы и системы программирования. – М.: Мир, 1990. – 320 с.

*Кравченко Вячеслав Александрович*, аспирант Восточно-Сибирского государственного университета технологий и управления, тел. +79644010435, e-mail: [krawyach@mail.ru](mailto:krawyach@mail.ru).

*Kravchenko Vyacheslav Aleksandrovich*, the post-graduate student of the East-Siberian state university of technologies and management.

УДК 681.51

©Л. М. Лукьянова, Л. Н. Федорченко

## СРЕДСТВА ФОРМАЛИЗАЦИИ ЦЕЛЕЙ И ПРОБЛЕМ СЛОЖНЫХ СИСТЕМ ПРОИЗВОДСТВЕННОЙ СФЕРЫ

Статья посвящена формализации формулировок целей и проблем производственных организационно-технических комплексов, которая способствует обоснованию системности решений по данным объектам.

**Ключевые слова:** производственная сфера, организационно-технический комплекс, модель формулировки проблемы/цели, формализованный язык проблем/целей.

L. M. Lukianova, L. N. Fedorchenko

## THE MEANS OF GOALS AND PROBLEMS FORMALIZATION IN INDUSTRIAL COMPLEXES

The means of goal and problem formalization in industrial organization-and-technical complexes are studied. Using the means is conduced to form system of decisions at problem situations.

**Keywords:** industry, organizational-and-technical complex, model of problem/goal, the formalized language of problems/goals.

### 1. Постановка задачи

Формирование решений по организационно-техническим комплексам производственной сферы (далее комплексы) сопряжено с трудностями, обусловленными сложностью данных объектов и недостаточной познаваемостью их свойств и поведения. Для лиц, принимающих решения (ЛПР), трудности возрастают в случае проблемных ситуаций, возникающих в комплексах и окружающей среде [1]. Неопределенность знаний ЛПР о данных объектах и большой вес плохо формализуемых субъективных факторов в методах коллективного формирования системы решений являются основными причинами позднего выявления несистемности решений по комплексам – во время и после их исполнения. С учетом фондоёмкости данных систем, создаваемой дорогостоящими техническими подсистемами, несвоевременная корректировка решений, не удовлетворяющих критерию системности, влечет издержки, ведет к частичному достижению целей и снижает эффективность функционирования комплексов [2, 3].

Типичными представителями данного подкласса организационных систем являются рыбохозяйственные комплексы (РХК). Большая фондоёмкость РХК создается рыбопромысловыми, транспортными, специализированными судами, а непредсказуемый спрос на рыбопродукцию и изменчивость состояний Мирового океана (сырьевой базы РХК) и увеличивают неопределенность знаний о РХК.

В работе [4] показано, что в условиях неопределенности вес субъективных моделей и неформализованных процедур особенно велик на начальных этапах процесса формирования решений по комплексу – на этапах анализа целей и проблем (под *целью* производственного комплекса понимается предвосхищение представляющего ценность результата его функционирования, под *проблемой* – противоречие между существующим и представляющим ценность будущим результатом), а системность целей является необходимым (однако не достаточным) условием системности решений.

С учетом некорректности результатов целеполагания и анализа целей, осуществляемых неформальными методами, обусловленности субъективных по своему характеру целей объективными проблемами и гипотезы о формулировании первых путем «отрицания» вторых [5], а также выявленной на большом корпусе формулировок соответствующих проблем и целей<sup>1</sup> корреляции их семантики, на этой же представительной выборке подтверждена возможность объективирования целей путем «отрицания» проблем.

Неформализованность средств анализа целей, приводящая к противоречивости и неполноте его результатов, делает актуальным в качестве первого шага решения поставленной задачи исследование моделей формулировок целей и проблем и построение на их основе

<sup>1</sup> Представительную выборку составили свыше пятисот проблем РХК и полутора тысяч целей комплексных межотраслевых направлений, покрывающих все виды деятельности данных объектов.



соответствующих формализованных языков.

## **2. Выбор класса языка целей/проблем**

Вследствие ориентированности на небольшие замкнутые миры и отсутствия прагматического и семантического компонентов формальные языки не позволяют решить задачу описания сложных целей/проблем комплексов.

Указанные недостатки частично преодолеваются в частично-формальных (формализованных) языках. Но в таких языках, как правило, фиксированы синтаксические структуры (и модели), которые для одних формулировок целей/проблем избыточны, а для других недостаточны. Кроме того, как показал анализ, семантический компонент данных языков не достаточно развит для адекватного описания проблем и целей, а прагматический – вовсе не реализован. Поэтому, в них не различаются основные и контекстные части формулировок целей/проблем. Все это делает целесообразной разработку формализованных языков целей (ФЯЦ) и проблем (ФЯП), базирующихся на более совершенных моделях их формулировок.

При разработке прагматического компонента ФЯЦ/ФЯП мы основывались на том, что целеполагание (образование и формулирование целей) и формулирование проблем – неформальные процедуры. В комплексах они могут осуществляться только ЛПР. Однако контроль прагматики пар <цель-подцель>, <проблема-подпроблема>, реализуемый на основе базовых знаний о комплексе, является формальной процедурой анализа целей и проблем, которую может осуществлять соответствующий анализатор.

При разработке семантического компонента ФЯЦ/ФЯП учитывалась необходимость: формализованного различения осмысленных целей/проблем среди синтаксически корректных, интерпретации отдельных знаков и их совокупностей в описаниях целей/проблем, выделения в последних целевых/проблемных и контекстных частей. Все это с учетом интеллектуального характера процессов формулирования проблем и целеполагания подтверждает ведущую роль прагматического и семантического компонентов ФЯП/ФЯЦ по отношению к синтаксическому и делает целесообразной целостную (не аддитивную) прагматико-семантико-синтаксическую формализацию цели/проблемы.

Такая целостная формализация цели/проблемы вызывает необходимость разработки ФЯЦ/ФЯП как ограниченного естественного языка (ЕЯ), относимого Н. Хомским к классу контекстно-зависимых (КЗ) языков [6]. С учетом известных трудностей разбора соответствующих описаний и современной классификации формальных языков и грамматик, КЗ-языки стремятся описывать обобщенными контекстно-свободными, атрибутивными, двухуровневыми грамматиками типа аффиксных или грамматик Ван Вейнгаардена [7]. Но они реализуют лишь синтаксический компонент языков и статическую семантику языка [13]. Этот факт и необходимость экспликации при анализе цели/проблемы, ее логической структуры позволило сделать заключение, о том, что ФЯЦ/ФЯП должен обеспечивать представление знаний о цели/проблеме.

В этой связи были проанализированы получившие распространение и подходящие для формализованного анализа целей/проблем классы языков представления знаний.

Как показал анализ указанной выше выборки формулировок целей/проблем, требованию целостной прагматико-семантико-синтаксической формализации в наибольшей степени удовлетворяют логические языки и модели, продукционные системы, семантические сети и сети фреймов. Вместе с тем логические языки при мощном механизме вывода обладают небольшой выразительной силой. Это затрудняет представление знаний о проблеме/цели и придает громоздкость их описаниям.

Выразительная мощность продукционных систем (при относительно простой реализации вывода) также недостаточна для представления знаний о проблеме/цели. Кроме того, для таких систем характерны трудности выработки стратегий управления выводом.

В сравнении с рассмотренными средствами семантические сети обладают лучшими выразительными возможностями, но число равноименных отношений, используемых при описании и анализе проблем/целей, требует большого числа процедур их обработки, что существенно усложняет реализацию сети.

Анализ результатов описания проблем/целей с помощью отмеченных средств показал эффективность (в том числе по критерию восприятия) «матрешечной» структуры фреймовых представлений [8]. Вместе с тем изменчивость проблем/целей в процессе функционирования

комплекса и разнородность частей их формулировок требуют разработки специфических фреймовых моделей с возможностями: настройки на предметную область и проблемную ситуацию; различения проблемных/целевых или контекстных частей формулировок; учета некоторых видов семантических отношений внутри слотов фрейма и между ними.

Теоретическим обоснованием фреймовых моделей цели и проблемы и соответствующих языков послужили следующие результаты лингвистических и логических исследований:

- язык цели [9] и семантико-синтаксические схемы формулировок целей в русском языке [10];
- универсальный семантический код [11], утилитарная положительная/отрицательная ценность [12], выражающая прагматическую связь проблем и целей [5];
- двухуровневые грамматики [7] и методы регуляризации формальных грамматик [13].

### 3. Фреймовая модель цели/проблемы

На русском языке цели производственных комплексов описываются инфинитивными предложениями [10], реализующими одну из двух структурных схем: минимальную – *Inf* или расширенную – *Inf < a >*, где *< a >* – расширение.

В целях комплексов намечается осуществление:

- средства («создать оборудование для производства пищевой продукции»). В этом случае ЕЯ-формулировка цели обычно включает в качестве контекстной части выделенный курсивом результат, реализуемый с помощью этого средства;

- результата («разработать новые виды копченой продукции, используя линию АЛ-10 на основе печи конвейерного типа»). В этом случае ЕЯ-формулировка цели обычно включает в качестве контекстной части выделенные курсивом средства осуществления результата.

Поэтому в формулировках целей присутствуют акции (действия), обычно обозначаемые глаголом в инфинитивной форме («разработать», «создать» и т.д.) или отглагольным существительным («для производства» и т.д.).

Это послужило обоснованием для выбора деятельностной модели цели «средства-результат» и использованию универсального семантического кода (УСК) [11] как информационного языка, кодирующего универсальную функциональную семантику языковых знаков, характеризующих действие. Путем расширения/сужения состава функциональных компонентов модель «средства-результат» легко настраивается на предметную область.

Обоснованием расширения/сужения фрейма «средства-результат» послужило следующее. Под «средствами» достижения цели и получаемыми с их помощью «результатами» в производственном комплексе естественно понимать ролевые (функционально различимые) компоненты деятельности. Теория УСК обосновывает систему синтаксических структур уровня универсальной функциональной семантики. В качестве базовой прагматико-семантико-синтаксической структуры модели цели была выбрана установленная данной теорией ядерная триада «агента-акция-объект» с определенными в [11] возможностями детализации ее компонентов.

Для решения поставленной задачи универсальные семантико-синтаксические структуры потребовали конкретизации. Анализ моделей действия, приведенных в [8], позволил определить семантико-синтаксический репертуар ролей фрейма цели производственного комплекса:

«агент – инструмент – способ – условие – объект».

Учитывая, что цель является моделью будущего результата или средства его осуществления, используемый в процессе целедостижения слот «акция» фрейма «средства – результат», при целеполагании заменяется на слот «способ». Более же широкую трактовку данного слота дает понятие «технология».

Целесообразно заменить и роль «инструмент» на более общую для производственной сферы роль «техника». Что касается условий, обычно выделяют условия целеобразования и целедостижения. Одно среди последних, – «место» целедостижения, включено в структуру фрейма цели.

Таким образом, с учетом приведенного в следующем разделе оператора целевой модальности, обеспечивающего различие между основной и контекстной частями формулировки цели, определена опорная прагматико-семантико-синтаксическая структура представления знаний о цели комплекса: «агент-техника-технология-место-объект», которая позволяет элиминировать функциональную (ролевую) многозначность в представлениях целей.

Поскольку деятельность по получению конечного результата комплекса детализируется последовательностью  $n$  действий по получению промежуточных результатов, в соответствии с теорией УСК допустимо расширение модели «средства–результат». Расширение фрейма цели в ширину состоит в левостороннем  $n$ -наращивании его опорной структуры, где  $n$  – «агнс–техника–технология». При этом разворачивается «горизонтальная» семантика фрейма. Расширение фрейма цели в глубину реализуется вложением опорных структур в отдельные его слоты. При этом осуществляется разворачивание «вертикальной» семантики фрейма.

Таким образом, построена фреймовая модель цели «средства–результат», иначе предложение–цель:

$$\langle \langle u^{ri} \rangle [\langle u^{rj} \rangle \dots] \rangle^2 \quad (1)$$

где

$u^{ri}$  ( $u^{rj}$ ) – номер или имя  $r_i$ -й ( $r_j$ -й) роли,  $n$  – число ролей,  $k \neq l$ ; парные угловые скобки–разделители, причём, внешние скобки выделяют предложение–цель, а внутренние разделяют его ролевые фразы.

Ограничив для простоты изложения число ролей в предложении–цели, отражающем минимальную ролевую дифференциацию компонентов операции «переработка» в РХК  $n = 7$ , и определив соответствующие им имена, имеем пример фреймовой модели цели:

$$\begin{aligned} &\langle \langle 1 - \text{агнс} \rangle \langle 2 - \text{технология управления} \rangle \\ &\langle 3 - \text{техника} \rangle \langle 4 - \text{исходный объект} \rangle \\ &\langle 5 - \text{технология производства} \rangle \\ &\langle 6 - \text{место} \rangle \langle 7 - \text{конечный объект} \rangle \end{aligned} \quad (2)$$

Аналогично построена фреймовая модель проблемы производственного комплекса, или предложение–проблема.

#### 4. Модель слота фрейма цели/проблемы

В результате эмпирического исследования формулировок целей/проблем производственных комплексов была определена общая модель слота фрейма цели/проблемы:

$$\text{роль\_базовый предмет} \{ \{ \_ \text{вид базового свойства} \} \{ \_ \text{базовое свойство} \} \}$$

где “\_” – разделитель, а фигурные скобки обозначают множественность заключенной в них части описания. Здесь квадратные скобки обозначают возможно пустой объект.

При этом под базовым предметом (БП) понимается простой (неделимый) предмет (денотат), определяемый лишь своим названием. Конкретизация свойств БП превращает его в производный предмет (ПП). Под базовым свойством (БС) понимается сторона ПП, обуславливающая его отличие или сходство с другими ПП. Для обеспечения адекватного описания и анализа целей/проблем выделены следующие виды БС (в скобках даны обозначения видов):

*характеристические* (СХ); данный вид свойств замещается БС семантической категорией, указывающей на признак ПП не прямо, а через отношение к другому предмету. Примером БС является (выделено курсивом): *судоремонтная инфраструктура*;

*функциональные* (СФ); данный вид свойств замещается БС как семантической категорией, указывающей на функцию ПП. Примеры соответствующих БС: консервирование, траление;

*именное* (СИ); данный вид свойств замещается БС как семантической категорией, персонализирующей ПП;

*физическое* (СЗ); данный вид свойств замещается БС как семантической категорией, создающей возможность измерения БС в порядковой или более «сильной» шкале. Примеры соответствующих БС: средний возраст, уровень профессиональной подготовки, загрузка инфраструктуры.

Для более глубокого описания физических свойств дополнительно введены: единицы измерения ( $M$ ) и значение ( $Z$ ) СЗ-свойства. Примеры  $M$  и  $Z$  (выделены курсивом): *средний возраст лет 50*.

<sup>2</sup> Квадратные скобки здесь и далее показывают факультативность или подразумеваемость заключенной в них части.

Таким образом, упрощенная базовая модель слота фрейма цели/проблемы имеет вид:

$$u^{ri} \text{ БП или } u^{ri} \text{ БП } [\{BC_k BC\}][\{CЗБСМЗ\}],$$

где  $u^{ri}$  – роль БП или ПП,  $BC_i$  – вид свойства,  $BC_i \in \{CX, C\Phi, СИ, CZ\}$ ,  $k = 1, 2, 3$ .

Базовая модель слота может быть расширена путем введения других видов свойств и помечена (вся или частично) оператором целевой  $G$  или проблемной  $H$  модальности, отражающим основную целевую или проблемную, утилитарно положительную или отрицательную ценность соответственно [12].

### 5. Грамматика и формализованный язык проблем/целей

Известно, что формальный язык относят к типу  $X$ , если существует хотя бы одна грамматика соответствующего типа, которая порождает данный язык. При разработке языка, в распознавании цепочек которого возникают трудности разбора, естественно стремление без потери его мощности перевести реализуемый язык из класса  $X$  в класс  $Y$ , например, из класса КЗ-языков в класс КС-языков, для которых отработаны алгоритмы разбора. Перевод осуществляется посредством эквивалентных преобразований грамматики исходного языка [13].

В данном контексте на основе предложенных фреймов проблемы и цели были разработаны соответствующие КЗ-грамматики с частичным учетом прагматики и семантики проблем и целей и осуществлен перевод их в эквивалентные двухуровневые КС-грамматики. Первый уровень определяет структуру предложения-проблемы/предложения-цели, а второй – структуру его ролевой фразы.

В качестве примера приведем грамматику и определяемый ею формализованный язык проблем. Грамматику второго уровня  $G_{mic}$  (для ролевой фразы предложения-проблемы) зададим так:

$$G_{mic} = \langle V_{Tmic}, V_{Nmic}, ОП_{mic}, P_{mic}, C_{mic} \rangle, \quad (3)$$

где

$V_{Tmic}$  – множество терминальных знаков;  $V_{Tmic} = \{\text{судовладелец, порты, ...}\}$ ;

$V_{Nmic}$  – множество нетерминальных знаков,  $V_{Nmic} = \{ОП_{mic}, \dots\}$ ;  $ОП_{mic}$  – цель грамматики;  $P_{mic}$  – множество продукций;

$C_{mic}$  – множество семантических условий при выполнении продукции,

$C_{mic} = \{C_{OCTy}, C_{OCTp}, C_{ATy}, C_{ATp}, C_{OoTy}, C_{OoTp}, C_{PMV}\}$ , где

$C_{OCTy}, C_{OCTp}, C_{ATy}, C_{ATp}, C_{OoTy}, C_{OoTp}, C_{PMV}$  – множество семантических условий по паре «агенса–место», «агенса–свойство», ..., «базовое свойство–единицы измерения\_значение» соответственно.

Граматику первого уровня  $G_{mac}$  (для предложения-проблемы) зададим так:

$$G_{mac} = \langle V_{Tmac}, V_{Nmac}, ОП_{mac}, P_{mac}, C_{mac} \rangle, \quad (4)$$

где  $V_{Tmac}$  – множество терминальных знаков;  $V_{Tmac} = \{<u^{ri}$  - фраза, ... $\}$ ,  $V_{Nmac}$  – множество нетерминальных знаков;  $ОП_{mac}$  – цель грамматики;  $P_{mac}$  – множество продукций;  $C_{mac}$  – множество семантических условий перехода после выполнения продукции,  $C_{mac} = \{C_{TcTa}, C_{TaTo}, C_{TcTaTo}\}$ , где  $C_{TcTa}, C_{TaTo}, C_{TcTaTo}$  – множество условий перехода.

Формализованный язык проблем зададим счетным множеством принадлежащих ему цепочек знаков  $prb_i$ , формируемых с помощью построенной выше двухуровневой грамматики  $G$  ( $G_{mac} \cup G_{mic}$ ):

$$\Phi ЯП(G) = \{prb_i\}.$$

Пример грамматического разбора двух формулировок проблем, заимствованный из [14] проиллюстрирован на рисунке (в скобках приведены семантические условия перехода).

Для обеспечения возможности настройки на предметную область грамматика ФЯП дополняется множеством функций  $\Psi$  изменения своего базиса. Дополнив грамматику (3) множеством  $\Psi_{min}$ , получаем грамматику  $G_{min1}$ :

$G_{min1} = \langle G_{min}, \Psi_{min} \rangle$ , где  $\Psi_{min} = \{\Psi_{Tmin}, \Psi_{Nmin}, \Psi_{Pmin}, \Psi_{Cmin}\}$ ,  $\Psi_{Tmin}, \Psi_{Nmin}, \Psi_{Pmin}, \Psi_{Cmin}$  – функции

изменения терминальных, нетерминальных знаков, правил вывода и семантических условий соответственно.

Таким же образом, т.е. путем дополнения грамматики (4) множеством  $\Psi_{\max}$ , получаем грамматику  $G_{\min 1}$ .

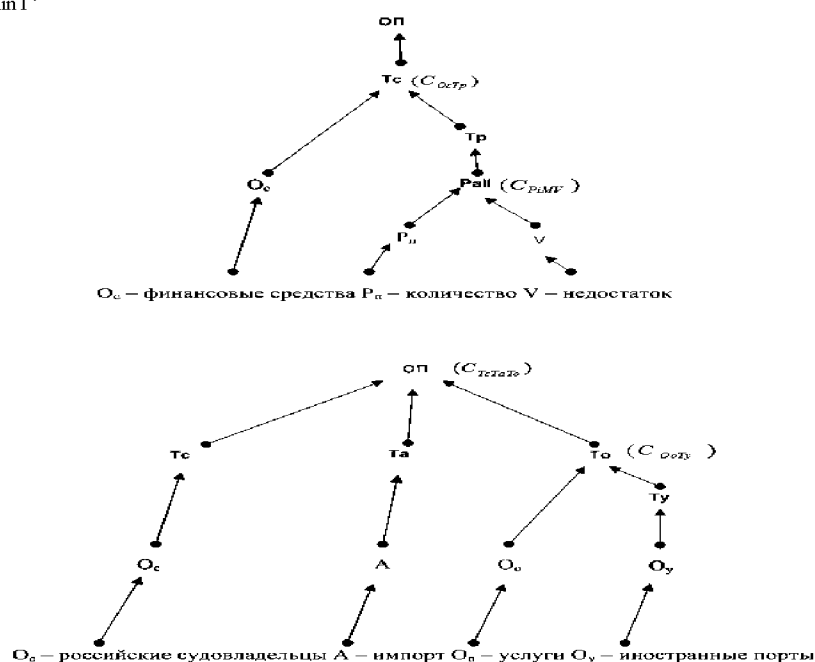


Рис.1. Пример разбора формализованного описания проблемы

### Заключение

Разработаны специфические фреймворки модели проблемы и цели производственного комплекса. На их основе построены двухуровневые формальные грамматики с частичным учетом прагматики, или обобщенной семантики (имена ролей, видов свойств и отношений: целевой и проблемной модальностей и фразовой прагматической согласованности), а также лексической семантики (базовые предмет и свойство, а для физических свойств – единицы измерения и значения) проблемы/цели. Рассмотрены задаваемые такими граммами формализованные языки проблем и целей. Указанные грамматики создают возможность создания эффективных анализаторов цепочек знаков, принадлежащих языкам, а сами языки позволяют адекватно описывать проблемы и цели комплексов.

Использование предложенных средств формализации способствует корректному формулированию, анализу и синтезу проблем и целей производственных комплексов, что является важнейшим условием обеспечения системности решений по данным объектам.

### Литература

1. Лукьянова Л. М. Трудности формирования системы решений по хозяйственному комплексу и основной путь их преодоления // Балтийский экономический журнал. 2010. №1(3). С. 84–91.
2. Исследование технологий анализа систем. Анализ проблемных ситуаций и путей их преодоления: отчет о НИР / КГТУ; рук. Л. М. Лукьянова. Калининград, 2000. 139 с. Деп.: (ВНИЦентр: №ГР 01.20.00 06420; инв. №02.20.00 04401).
3. Исследование технологий анализа систем. Проблемный анализ рыбохозяйственного комплекса РФ: отчет о НИР / КГТУ; рук. Л. М. Лукьянова. Калининград, 2009. 56 с. Деп.: (ВНИЦентр: №ГР 01200006420; инв. №02200903520).
4. Лукьянова Л. М. Результаты развития методологии формирования решений по организационно-техническим комплексам // Известия КГТУ. 2010. №19. С. 36–44.
5. Тамбовцев В. Л. Логическая структура прогнозирования и целеполагания // Целевая стадия планирования и проблемы принятия социально-экономических решений. М.: ЦЭМИ АН СССР, 1972. С. 119–131.
6. Хомский Н. Формальные свойства грамматик / Кибернетич. сб. Новая серия. Вып. 6. М.: Мир, 1969. С. 65–98.
7. Wijngaarden A. van. The generative power of two-level grammars // Automata, Languages and Programming. pp. 9-16; Lecture Notes // Computer Science. №14, / ed. J. Loeckx, Springer-Verlag, Berlin, 1974.

8. Минский М. Фреймы для представления знаний. М.: Энергия, 1979. 192 с.
9. Логический анализ языка. Избранное. 1988-1995 / редкол.: Н. Д. Арутюнова, Н. Ф. Спиридонова; РАН: Ин-т языкознания. М.: Индрик, 2003. 696 с.
10. Современный русский язык: учебник для филол. спец. вузов / В.А. Белошапкова, Е.А. Брызгунова, Е.А. Земская и др. 3-е изд., испр. и доп. М.: Азбуковник, 1999. 928 с.
11. Мартынов В. В. УСК – язык представления знаний и эвристического поиска. М.: Науч. совет по комплексной проблеме «Кибернетика», 1981. 87 с.
12. Ивин А. А. Основания логики оценок. М.: Изд-во МГУ, 1970. 230 с.
13. Федорченко Л. Н. Книга – Регуляризация контекстно-свободных грамматик. Saarbrücken, Germany: LAP LAMBERT Academic Publishing GmbH & Co. KG Dudweiler Landstr. 99, 66123, 2011. 180 с.
14. Зинько Д. П., Лукьянова Л. М. Формализованный язык проблем организационно-технических комплексов // Вестник Калининградской академии информатологии МАИ: сб. науч. трудов. Калининград: КГТУ, 2011. С. 92–103.

*Лукьянова Людмила Михайловна* – д-р техн. наук, проф. каф. систем управления и вычислительной техники ФГБОУ ВПО «Калининградский государственный технический университет», академик Международной академии информатизации (МАИ), тел. +7 (4012) 995-942. E-mail: [llm\\_llm@mail.ru](mailto:llm_llm@mail.ru)

*Федорченко Людмила Николаевна* – канд. техн. наук, старший научный сотрудник лаборатории прикладной информатики Федерального государственного бюджетного учреждения науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Тел. +7 (812) 328-1919. E-mail: [LNF@iias.spb.su](mailto:LNF@iias.spb.su)

*Lukianova Lyudmila Mikhailovna* – doctor of technical science, professor of computer science department in FSSO HPS «Kaliningrad State Technical University».

*Fedorchenko Ludmila Nikolaevna* – Senior Research Associate, PhD, St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS)

E-mail: [LNF@iias.spb.su](mailto:LNF@iias.spb.su).

УДК 620.9 : 658.011.56

©И.У. Сахабетдинов

## **ЗАДАЧИ КОМПЛЕКСНОГО МОНИТОРИНГА В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ ДИСПЕТЧЕРСКОГО УПРАВЛЕНИЯ ЭНЕРГЕТИЧЕСКИМИ ОБЪЕКТАМИ**

Рассматриваются принципы создания современных автоматизированных систем диспетчерского управления, а также функции входящих в их состав подсистем мониторинга различного назначения. Оцениваются возможности интеграции перспективных технологий в существующую структуру управления энергетическими объектами.

**Ключевые слова:** автоматизированные системы диспетчерского управления, системы мониторинга, оперативно-информационные комплексы.

*I.U. Sakhabetdinov*

## **CHALLENGES OF INTEGRATED MONITORING IN AUTOMATED SYSTEMS DISPATCHING ENERGY FACILITIES**

The principles of creation of modern automated systems, supervisory control, as well as the functions of their component sub-systems monitoring for various purposes. The possibility of integrating emerging technologies into the existing management structure of energy facilities.

**Keywords:** automated systems for supervisory control, monitoring, operational and information systems.

### **Введение**

Для эффективного и надежного функционирования электроэнергетических объектов (энергообъединений, энергосистем, сетевых и генерирующих предприятий) необходимо создание и внедрение современных систем автоматизации оперативно-диспетчерского и технологического управления.

Последние годы отмечены реформированием российской энергетики. Как показывает международный опыт, переход к рыночным принципам функционирования в электроэнергетике повышает эффективность работы и увеличивает размер прибыли, искореняя затратные тенденции, снижая издержки, что ведет к существенному росту заинтересованности инвесторов во вложении средств в этот сектор.

В соответствии с Положением о Центральной диспетчерской службе (ЦДС) Федеральной сетевой компании (ФСК), основными задачами ЦДС являются: мониторинг состояния Единой национальной электрической сети (ЕНЭС) и оперативный контроль проводимых ремонтных и ремонтно-восстановительных работ, организация оперативных действий по оптимизации режима работы, локализация аварий и восстановление объектов ЕНЭС в пределах полномочий ФСК, организация работ по оптимизации электрических режимов с целью уменьшения расхода электроэнергии (мощности) на ее транспорт по ЕНЭС, организация и контроль работы оперативного персонала в филиалах ОАО «ФСК ЕЭС». В соответствии с этим, основными задачами, которые должны оперативно решаться в ЦДС ФСК, являются: сбор информации и анализ оперативной обстановки на объектах ФСК, определение места и характера повреждений линий электропередачи и оборудования подстанций, определение причин нарушения режима работы и аварийных отключений оборудования, участие в разработке и контроле мероприятий, предотвращающих подобные ситуации, оперативный контроль аварийно-восстановительных работ, включая использование аварийного запаса, проработка оперативных заявок на вывод в ремонт оборудования электросетевого хозяйства ФСК, организация взаимодействия с Системным Оператором (СО), контроль и оценка влияния плановых и оперативных решений СО по режимам работы электрических сетей ФСК на надежность и экономику объектов. К этим задачам оперативного управления примыкает задача подготовки оперативного персонала подразделений ФСК.

Достижение этих целей невозможно без автоматизированной современной инфраструктуры, включающей перспективные программно-технические средства организации производства и автоматизации энергетических систем.

К основным оперативным функциям диспетчерского управления в ФСК относится осуществление мониторинга состояния ЕНЭС, мониторинга потерь электроэнергии. Само понятие «мониторинг» подразумевает такие действия, как сбор и оценка информации, контроль объекта (или

процесса) и прогнозирование его поведения. Все эти составляющие понятия «мониторинг» могут реализовать специализированные подсистемы, входящие в состав любой современной автоматизированной системы диспетчерского управления (АСДУ).

Основой автоматизированных систем диспетчерского управления являются оперативно-информационные комплексы (ОИК), дополненные комплексами для решения технологических задач. Современные требования к представлению и использованию информации в интегрированных информационных системах делают целесообразным использование таких понятий и технологий, как активно-адаптивные сети (Smart Grid) и СИМ-модели.

В России термин «Smart Grid» трактуют как «интеллектуальную сеть энергетики», «интеллектуальную электроэнергетическую систему», «активно-адаптивную сеть энергетики» [2]. Представляется целесообразным рассматривать понятие «интеллектуальная электроэнергетическая система» как совокупность технологий и стандартов, объединяющих объекты генерации электроэнергии (крупные и малые), оборудование электрических сетей (магистральных, распределительных) и энергопринимающие устройства потребителей в единую информационную систему, в которой осуществляются мониторинг и управление в реальном времени работой всех участников производства, передачи, сбыта и потребления электроэнергии, в том числе в аварийных режимах. Технология Smart Grid представляет собой систему, оптимизирующую энергозатраты и позволяющую перераспределять электроэнергию. При традиционном распределении электричества ток по проводам поступает от станции к потребителю и подается в соответствии с заранее заданным уровнем напряжения и сопротивления. Внедрение "интеллектуальных" сетей позволит автоматически регулировать подачу электроэнергии в зависимости от снижения или увеличения режима потребления.

Common Information Model (CIM) – общая модель информации – «абстрактная модель, которая все множество элементов электроэнергетической системы представляет стандартным образом в виде описания объектов, их свойств и связей между ними. Такое единое описание позволяет осуществлять интеграцию различных приложений, выполненных независимыми изготовителями» (МЭК-61970-301).

СИМ представляет собой некоторую концептуальную модель для описания различных предметов (объектов) окружающего мира, используя объектно-ориентированную терминологию. Если до последних лет понятия объектно-ориентированной технологии относились к языкам программирования (C++, Java и др.), то СИМ расширяет эти понятия до описания информационного пространства, сознательно используя такую терминологию объектно-ориентированного программирования как классы, свойства, методы и ассоциации. По существу СИМ представляет собой информационную модель, задачей которой является единое унифицированное представление структур данных, независимо от источника происхождения данных и целей их использования [3].

Основными целями работ при создании современных отечественных ОИК, как правило, являются:

- 1) определение основных направлений развития автоматических систем оперативно-диспетчерского управления энергообъединениями и электросетями в условиях реформирования отечественной энергетики; определение основных функций комплексного мониторинга состояния ЕЭС/ЕНЭС; определение основных функций оперативного мониторинга коммутационного состояния электрической сети; определение общей структуры оперативного контроля электрического режима ЕЭС/ЕНЭС; определение основных функций мониторинга и анализа оперативных ремонтных заявок; определение понятия и основных функций мониторинга готовности оперативно-персонала;

- 2) обобщение опыта разработки оперативно-информационных комплексов, выработка рекомендаций по использованию различных типов отечественных ОИК, определение перспектив развития информационной архитектуры ОИК с учетом специфики реструктуризации электроэнергетики;

- 3) разработка методологии создания обобщенных информационных моделей (СИМ-систем), как интеграционного механизма для информационных взаимодействий в системах управления ЕНЭС; разработка мероприятий по адаптации СИМ-систем к современным отечественным условиям;

- 4) разработка методологии использования комплекса экспертных систем в ОИК энергообъединений и сетевых компаний, определение функций экспертных систем, входящих в этот комплекс;

- 5) разработка экспертной системы-советчика для проработки оперативных заявок;



- 6) разработка экспертной системы для анализа топологии электрической сети;
- 7) внедрение разработанных комплексов программ и рекомендаций в практику работы систем управления энергосистемами и электросетями.

Для достижения этих целей необходимо решить следующие основные задачи:

- 1) разработка методологии и алгоритмических принципов осуществления мониторинга состояния ЕЭС и ЕНЭС и функций мониторинга готовности оперативного персонала;
- 2) разработка обобщенной архитектуры комплекса ОИК АСДУ и формальных алгоритмов обработки информации с целью унификации технических требований и возможности сертификации оперативно-информационных комплексов;
- 3) анализ возможностей и подготовка предложений по адаптации обобщенных информационных моделей к условиям отечественных систем управления ЕНЭС;
- 4) разработка методологии использования, структуры и функций комплекса экспертных систем для оперативно-диспетчерского управления в ЕЭС и ЕНЭС; разработка алгоритмических принципов построения экспертных систем для режимной проработки ремонтных заявок и для анализа топологии электрических сетей;
- 5) обобщение опыта разработки аналогичных программных комплексов и систем в России и за рубежом [4].

## **1. Автоматизированное диспетчерское управление в условиях реструктуризации российской энергетики; мониторинг состояния ЕЭС и ЕНЭС**

*1.1. Совершенствование систем автоматизированного диспетчерского управления в нормальных и аварийных режимах; существующие технологии в энергетике.*

Управление режимами ЕЭС России, включая все ее звенья, подчинено единой цели – обеспечению надежной и экономичной работы при рациональном расходовании энергоресурсов и при бесперебойном энергоснабжении потребителей электроэнергией требуемого качества. Реализация этих целей осуществляется автоматизированной системой диспетчерского управления, которая состоит из средств вычислительной техники, связи, телемеханики, систем автоматики и комплексов программного обеспечения [5].

АСДУ представляет собой иерархически построенную человеко-машинную систему, обеспечивающую по всей территории, охватываемой электрическими сетями, сбор, преобразование, передачу, переработку и отображение информации о состоянии и режиме системы, формирование на основе собранной информации, передачу и реализацию управляющих команд с целью выполнения системой (за счет располагаемых средств) функций надежного и экономичного снабжения электрической и тепловой энергией требуемого качества всех ее потребителей [6].

В составе АСДУ ЕЭС на всех уровнях ее иерархии созданы и эксплуатируются:

- системы автоматического управления режимами – системы релейной защиты, противоаварийной автоматики и автоматического управления нормальными режимами по частоте и активной мощности (АРЧМ);
- оперативно-информационные и управляющие комплексы (ОИУК), обеспечивающие в реальном времени дежурного диспетчера информацией о текущем режиме, управление диспетчерским щитом, ведение суточной диспетчерской ведомости и пр.;
- системы оперативного управления внутри суточного периода (советчик диспетчера), обеспечивающие внутрисуточную коррекцию режима по активной мощности и напряжению, оперативную оценку работоспособности;
- системы краткосрочного (сутки, неделя) и долгосрочного (месяц, квартал, год) планирования энергетических и электрических режимов;
- системы автоматизации коммерческого учета и контроля электроэнергии и мощности (АСКУЭ);
- экспертные системы для информационной помощи оперативному персоналу;
- диспетчерские тренажеры.

В реализации непрерывного оперативно-технологического управления режимами работы ЕЭС России большую роль играет оперативно-информационный и управляющий комплекс, являющийся важной подсистемой автоматизированной системы Системного Оператора (АС СО). Этот комплекс, основанный на использовании текущей телеинформации и результатов решения задачи оценки состояния, имеет два вида назначений: текущее обслуживание диспетчеров ЕЭС и ее звеньев, а также формирование массивов информации для задач оперативного управления и пла-

нирования режимов.

К настоящему времени разработаны и используются множество ОИК, причем компании-разработчики выступают как независимые структуры, и по этой причине системы не унифицированы и даже не всегда снабжены протоколами обмена данными. Поэтому важнейшей задачей в этом направлении является проблема унификации, отбора лучших систем для разных уровней иерархии управления. Унификация должна касаться многих аспектов – как унификации информационного описания объектов, так и использования унифицированных баз данных, унифицированных средств отображения информации (в том числе на основе типовых технологий Internet), используемых операционных систем, «клиент-серверных» технологий, сетевых протоколов, средств доступа к данным и человеко-машинного интерфейса. Роль задач ОИК постоянно возрастает по мере усложнения условий функционирования электроэнергетики. Если несколько лет назад ОИК, в основном, применялись как средства приема, обработки и отображения телемеханической информации, то сегодня перед ОИК ставятся все новые и новые задачи по управлению электропотреблением, обеспечению надежности и экономичности энергоснабжения, обеспечению функционирования рынка и т.п.

Важную роль в этих задачах помощи диспетчеру начинают играть экспертные системы. Экспертные системы могут использоваться: как консультанты диспетчера (например, в задаче ремонтных заявок), как советчики диспетчера (например, по управлению оборудованием), для анализа ситуаций (оперативная диагностика «нештатных» ситуаций), для диспетчерского мониторинга (интеллектуальный анализ топологии электросетей) и как тренажеры.

Значительную роль в рыночных условиях занимает автоматизированная система коммерческого учета электроэнергии (АСКУЭ), которая также является подсистемой АС СО. АСКУЭ предназначена для осуществления мониторинга и коммерческого учета балансов мощности и энергии субъектов рынка и обеспечивает получение данных о средних значениях мощности и об учтенной электроэнергии по зонам суток за каждые календарные сутки и накопительно за неделю, месяц, год. Приборы коммерческого учета устанавливаются на энергообъектах и обеспечивают передачу накопленной информации по каналам связи на ближайший по иерархии диспетчерский пункт. Информация поступает в компьютер, обрабатывается (агрегируется) и передается на следующий уровень управления, вплоть до верхнего, а также в информационно-вычислительную систему энергосбыта.

Противоаварийное управление (ПА) является средством повышения надежности работы ЕЭС РФ. В России уже создана развитая система противоаварийного управления, обеспечивающая значительное повышение надежности энергоснабжения потребителей и живучести ЕЭС России. Противоаварийная автоматика предотвращает или прекращает нарушение устойчивости параллельной работы энергосистем, что значительно сокращает число и вредные последствия наиболее тяжелых системных аварий [7].

Помимо программных средств, определяющую роль в разработке и функционировании современных АСДУ играют технические средства. На текущий момент на некоторых энергопредприятиях России ещё работают морально и физически устаревшие устройства телемеханики и передачи информации, к тому же объем телемеханики имеется в недостаточном объеме, не обеспечивающем требуемой наблюдаемости сети. Поэтому ЦДУ ЕЭС отмечает необходимость опережающего развития средств телемеханики, связи и систем приема-передачи технологической информации. В перспективе должна быть создана Единая сеть связи на базе широкого внедрения современных коммутационных узлов, волоконно-оптических линий связи, использования аппаратуры синхронной цифровой иерархии, а также системы спутниковой связи [8].

### *1.2. Мониторинг состояния ЕНЭС.*

Система мониторинга состояния является подсистемой общей системы АСДУ. В узком смысле мониторинг состояния включает средства измерения (автоматические или для ручного применения) необходимых параметров и средства для их обработки, а также средства хранения и отображения. При более углубленной обработке информации осуществляется проверка некоторых установленных критериев для этих показателей. На основе полученных данных принимаются решения в отношении состояния (на той или иной временной стадии) для того или иного объекта [9].

С учетом этого, средства и инструменты мониторинга состояния не должны без необходимости обособляться, а, наоборот, по возможности, должны интегрироваться в состав более общих технических, программных, информационных средств. Это тем более необходимо, так как зачастую в подсистемах АСДУ используется одна и та же информация. Отсюда следует необходимость инте-

грации с соблюдением наиболее жестких требований из всех охватываемых интеграцией функций. В частности, на любом объекте система мониторинга должна быть погружена в соответствующие общие системы автоматизации управления, а при осуществлении её на неавтоматизированном объекте или объекте с устаревшей системой автоматизации, она должна разрабатываться как подсистема будущей системы управления. Подходы к реализации системы должны быть основаны на современных информационных технологиях с максимальной их унификацией, соблюдением общих стандартов передачи, обработки, отображения информации, операционных сред, баз данных, метрологических стандартов, общих стандартов программирования и др. Особенно важна, в частности, унификация принципов обмена информацией между программными системами, между объектами и субъектами.

Средством обеспечения диспетчерского персонала сетевого оператора и компаний электрической сети необходимой информацией о надежности сети является комплексный мониторинг, включающий следующие взаимосвязанные компоненты (наименования условные):

- прогноз-мониторинг состояния ЕНЭС,
- ретроспективный мониторинг состояния ЕНЭС,
- оперативный мониторинг состояния ЕНЭС,
- мониторинг потерь электроэнергии.

Ретроспективный мониторинг должен содержать:

- ретроспективный анализ системных аварий, аварийности линий и оборудования подстанций,
- анализ многолетних архивов графиков активных и реактивных нагрузок узлов сети ЕНЭС.

Анализ аварийности оборудования нужно выполнять с учетом даты ввода в эксплуатацию, дат ремонтов, сведений по модернизации оборудования, сведений о месте установки (присоединения) оборудования, о климатических условиях, о максимальном токе короткого замыкания в месте установки, от количества отключений токов КЗ, от режима работы, предшествующего повреждению. Анализ повреждаемости элементов воздушных линий электропередачи (ВЛ) необходимо выполнять с учетом паспортных данных и данных о виде и характере повреждений, проявившихся в предыдущие периоды эксплуатации. Должны учитываться имевшие место факты вандализма, фактические климатические воздействия и расчетные климатические нагрузки, принятые при проектировании ВЛ.

К ретроспективному мониторингу может быть отнесена система фиксации и анализа (расследования) произошедших технологических нарушений, которая предназначена для установления причин, последствий и ответственности участников. Система должна охватывать всех субъектов и инфраструктуру рынка. Процедура расследования нарушений должна быть регламентированной и должна координироваться уполномоченным органом.

Прогноз-мониторинг состояния ЕНЭС должен включать прогнозирование развития ЕНЭС, прогнозирование нагрузок, определение и коррекцию набора типовых ремонтных и аварийных режимов.

Оперативный мониторинг состояния электросети включает, по крайней мере, следующие компоненты:

- оперативный мониторинг топологического состояния и режимов работы электрической сети;
- мониторинг и анализ оперативных заявок на вывод из работы оборудования подстанций и линий,
- мониторинг качества работы оперативного персонала.

Мониторинг топологического состояния (МТС) электросети является одним из основополагающих компонентов оперативного мониторинга состояния ЕНЭС. Основными функциями МТС являются:

- определение топологического состояния сети, выявление и отображение для диспетчерского персонала сетевых событий, влияющих на надежность электросети; такими событиями являются отключения (подключения) ЛЭП, силовых трансформаторов, генераторов, реакторов, разделение схем подстанций, отделение энергообъектов и районов сети от ЕНЭС;
- использование результатов анализа топологии сети для определения контролируемых пределов режимных параметров (в основном, потоков активной мощности ЛЭП и в контролируемых сечениях); эти пределы задают область надежной работы сети;
- представление результатов анализа топологии сети для использования режимными технологическими программными комплексами – системами оперативного управления (типа «советчик диспетчера»), системами АСКУЭ и др.

Ремонтные заявки являются важной составляющей информации для оперативного управления в энергосистемах и электросетях. Ремонт необходимого элемента оборудования в соответствии с принятой в отечественном диспетчерском управлении технологической практикой заблаговременно оформляется специальным документом – ремонтной заявкой, содержащим ряд атрибутов, включая:

- наименование элемента оборудования,
- вид операции (вывод в ремонт, ввод в работу),
- вид заявки (плановая, «внеплановая», аварийная),
- время начала и окончания работ по заявке.

Заявки должны быть проработаны на тех энергопредприятиях, в чьем управлении и ведении находится соответствующий элемент оборудования, причем эта проработка осуществляется оперативными службами (служба электрических режимов, служба релейной защиты и автоматики, диспетчерскими службами) диспетчерских управлений в соответствии с иерархией оперативного диспетчерского управления.

Мониторинг и анализ информации по оперативным ремонтным заявкам является важной составляющей оперативного мониторинга надежности в ЕНЭС/ЕЭС. Можно выделить следующие функции мониторинга по заявкам:

- автоматизированная проработка оперативных заявок,
- мониторинг прохождения заявок,
- выявление «нештатных» ситуаций в электросети с использованием информации по открытым заявкам,
- «диспетчерская» автоматизированная проработка заявок,
- автоматическая проверка выполнения режимных условий для открытых ремонтных заявок,
- достоверизация и пополнение информации о коммутационном состоянии электросети на основе телесигналов и информации по открытым заявкам.

Оперативно-диспетчерский и эксплуатационный персонал является составной частью человеко-машинной системы управления на электростанциях, подстанциях и в электрических сетях, на диспетчерских пунктах СО, надежность его действий является определяющим условием эффективности управления.

Эффективность профессиональной деятельности оперативного персонала рассматривается в двух аспектах:

- готовность выполнения персоналом его функций в составе упомянутой системы, в том числе при осуществлении необходимых действий,
- безопасность персонала, при выполнении этих функций.

Мониторинг готовности для оперативного персонала включает, кроме упомянутого, регистрацию и анализ действий при его функционировании, а также анализ результатов противоаварийных тренировок, в том числе на тренажерах.

*1.3. Перспектива развития технических и информационных средств мониторинга состояния; автоматизация подстанций.*

*1.3.1. Обеспечение измерений.*

В области измерительной аппаратуры необходима замена недостаточно точных измерительных трансформаторов тока и напряжения на энергообъектах на современные трансформаторы с более высоким классом точности, в том числе применение:

- элегазовых трансформаторов тока на напряжение 110 кВ и выше с высоким классом точности (в т.ч. 0,2), обладающих повышенной надежностью и пожаробезопасностью;
- емкостных трансформаторов напряжения с повышенным классом точности (до 0,2);
- комбинированных трансформаторов тока и напряжения в одном корпусе.

В области средств телемеханики и связи при реконструкции систем сбора и передачи информации основными требованиями являются:

- использование цифровых каналов связи;
- использование на энергообъектах специализированных цифровых телемеханических комплексов (а для электростанций и крупных подстанций – ППК АСУТП), которые предоставляют возможность ввода аналоговой информации непосредственно от трансформаторов тока (ТТ) и напряжения (ТН), обладают значительным вычислительным ресурсом и более совершенным программным обеспечением;
- замена центральных приемо-передающих станций на программно-технические комплексы,

использующие современные достижения в области вычислительной техники и цифровых систем связи,

- внедрение указанных ПТК в оперативно-информационные комплексы современных SCADA-систем.

Аппаратура телемеханики нового поколения должна обеспечивать интеллектуальную обработку информации, передаваемой на ДП (достоверизация, суммирование ТИ, формирование обобщенных ТС и др.), организацию отображения информации на рабочем месте оперативного персонала энергообъекта, дистанционное управление коммутационным оборудованием с рабочего места оперативного персонала энергообъекта, сбор и передачу ретроспективной информации (дискретные и аналоговые регистраторы событий, интегральные значения параметров) [10].

Передача телеинформации должна осуществляться по нескольким направлениям с возможностью использования различных протоколов обмена и состава передаваемой информации, реализации адаптивного способа передачи, передачи телеинформации на пункты управления не только по выделенным, но и по коммутируемым каналам связи (для передачи ретроспективной информации). Перспективными являются цифровые каналы, резервированные по разным трассам, при этом аналоговые каналы телемеханики подлежат замене на цифровые. Скорость передачи по каналам определяется в зависимости от объемов трафика и ограничений, накладываемых устройствами передачи и обработки информации, коэффициент готовности каналов не ниже 0,999, время восстановления не более 5 минут. Передача информации регистраторов аварийных событий должна осуществляться в соответствии с требованиями к каналам передачи технологической информации, регистраторы должны быть масштабируемыми по видам интерфейсов для сопряжения с каналами передачи данных.

### *1.3.2. Автоматизация подстанций.*

Решение задачи мониторинга отказов наиболее полно обеспечивается при высокой автоматизации подстанций.

Автоматизация управления с интеграцией систем измерения, автоматики, защиты, диагностики и управления оборудованием в магистральных и распределительных электрических сетях обеспечивает повышение эффективности функционирования объекта в целом (подстанции, сетевого района, предприятия магистральных электрических сетей) в нормальных и аварийных режимах, снижение аварийных ущербов и потерь, снижение эксплуатационных затрат и затрат на ремонт основного и вспомогательного оборудования.

Основными направлениями развития в области автоматизации электросетевых объектов должны стать: разработка и внедрение системы мониторинга на основе SCADA-систем различного уровня, внедрение новых типов сетевого оборудования, предназначенных для работы в составе полностью автоматизированных технологических комплексов, внедрение системных микропроцессорных устройств измерений, защиты, автоматики и управления в составе АСУ ТП, внедрение новых подсистем контроля и мониторинга, обеспечивающих решение задач оперативного получения всесторонней объективной информации о выполнении всеми субъектами рынка энергии и мощности договорных обязательств в нормальных и аварийных режимах работы энергообъединений, жесткий контроль выполнения условий технической и программной совместимости всех систем управления ЕНЭС, в том числе при смене поколений вычислительных средств и при использовании устройств иностранного производства. Необходимо внедрение единых информационно-диагностических систем, использующих интеллектуальные способы и оценки.

АСУ ТП подстанций должны создаваться таким образом, чтобы их функционирование было инвариантным по отношению оперативно-технологическому и организационно-экономическому структурам управления.

### *1.3.3. Информационное и программное обеспечение.*

Фиксируемые данные по любому элементу должны включать: паспортные данные (если имеются) и информацию о «жизненном пути» элемента, в том числе место и время ввода в эксплуатацию, данные по техобслуживанию и ремонтам, сведения по модернизации, информацию по отказам. Должны использоваться распределенные базы данных с возможностью размещения их разделов на объектах и структурных единицах данного субъекта. Должны быть определены информационные интерфейсы (протоколы обмена) между всеми взаимодействующими подсистемами системы мониторинга и дисциплина внешних обменов данными.

Специальное программное обеспечение должно выполнять иерархическую обработку данных с получением установленных показателей для каждого контролируемого элемента каждого субъек-

та, в том числе обобщенных показателей для субъектов и ЭЭС – крупных технологических единиц, а также в целом по ЭЭС и отрасли. Должна быть установлена рациональная степень унификации программного обеспечения с учетом групп однородных субъектов и объектов.

Системотехническое обеспечение включает в себя технические и стандартные программные средства получения, передачи, обработки и представления (отображения) информации, а также средства управления процессами мониторинга. Ко всем этим средствам с учетом их функций должны быть предъявлены технические требования исходя из необходимости для каждого субъекта и в отрасли в целом распределенной организационно-технической автоматизированной системы, реализующей мониторинг надежности. При этом должна быть обеспечена достаточная гибкость и установлена рациональная степень унификации системотехнических решений, в том числе по группам однородных субъектов и объектов.

## 2. Основные функции оперативно-информационных комплексов

Оперативно-информационные комплексы автоматизированных систем диспетчерского управления в энергосистемах (ОИК АСДУ) являются наиболее важными средствами для обеспечения оперативно-диспетчерского и руководящего персонала энергообъединений, энергосистем, энергопредприятий компьютерной поддержкой при принятии оперативных решений по управлению. Широко применяемые во всем мире комплексы ОИК (или SCADA – системы мониторинга, управления и сбора данных) внедрены в тысячах энергетических предприятий.

Основные функции ОИК АСДУ для всех уровней управления определены следующим образом:

- ввод в ОИК текущей оперативной информации;
- первичная и вторичная обработка оперативной информации;
- архивирование;
- обеспечение диалога пользователей с ОИК;
- отображение информации пользователям ОИК;
- обеспечение надежности функционирования ОИК.

Эти функции являлись и остаются до настоящего времени стандартными функциями современных ОИК, при этом в любом комплексе можно выделить две обязательные составляющие:

- система обработки информации («центральные» серверы);
- система человеко-машинного интерфейса (автоматизированные рабочие места – АРМ).

В начале работ по созданию ОИК были определены основные функции системы обработки телеинформации: фиксация изменений телесигналов, масштабирование телеизмерений, усреднение измерений с разными периодами, сравнение измеренных параметров с технологическими пределами и плановыми значениями, достоверизация измеренных параметров, архивирование оперативной информации, «дорасчет» параметров и сигналов.

С помощью ОИУК диспетчерский персонал ЦДУ, ОДУ, ЦДС осуществляет контроль за текущим состоянием управляемой ЭЭС (схемой, режимами, средствами управления), ретроспективный анализ происшедших событий, оценку перспективных режимов. На основании информации о текущем и перспективном состоянии ЭЭС, на графике нагрузки, плане проведения ремонтных работ по оперативным заявкам с учетом указаний и рекомендаций диспетчерских инструкций и директивных материалов диспетчерский персонал обеспечивает: выработку воздействий на управляемые объекты (регулирование режима ЭЭС по активной и реактивной мощности, включая регулирование графиков нагрузки электростанций); вывод оборудования, а также средств автоматического и оперативного управления в ремонт и ввод их в работу после ремонта; ввод в работу нового оборудования и средств управления; изменение схемы коммуникации контролируемой сети; ликвидацию аварийных ситуаций и восстановление нормального режима работы; ведение оперативной отчетности и передачу оперативной информации.

Управляющие воздействия передаются диспетчерским персоналом ЦДУ, ОДУ или ЦДС на оперативно подчиненные объекты через диспетчерский персонал этих объектов либо непосредственно на АСУ ТП и системы автоматического регулирования и управления энергообъектами с помощью устройств телеуправления. Управляющими воздействиями обеспечивается изменение схемы электрической сети или состава оборудования электростанций и подстанций; алгоритмов и параметров настройки средств автоматического и оперативного управления, устройств автоматики; нагрузки агрегатов электростанций; нагрузки потребителей; напряжений в контрольных точках электрической сети.

### 2.1. Информация, используемая при управлении.

Для успешного решения задач оперативно-диспетчерского управления диспетчер должен располагать необходимой, достаточно достоверной информацией. Качественно новые аспекты информационного обеспечения АСДУ связаны с привлечением и использованием информации:

- о прогнозе метеорологической обстановки для повышения точности прогнозирования нагрузок и вероятностей отказов;
- о маневренных характеристиках агрегатов и электростанций, используемых при расчете их располагаемой и рабочей мощности и состава работающего и резервного оборудования;
- об отказах основного оборудования ЕЭС с целью расчета и прогнозирования его показателей надежности;
- о качестве топлива, поставляемого тепловым электрическим станциям (ТЭС);
- о состоянии основного оборудования (генерирующего оборудования, ЭП, трансформаторов, атомных реакторов и др.), определяемом в процессе диагностики для принятия с необходимой своевременностью решения о времени его вывода в ремонт;
- о фактически обеспечиваемой надежности электро- и теплоснабжения потребителей для выбора оптимальных путей повышения надежности;
- о прогнозе притока воды в водохранилищах гидроэлектростанций (ГЭС) с целью оптимизации выработки электроэнергии на ГЭС.

Необходимая информация либо поступает извне, либо вырабатывается (извлекается) в процессе управления [11].

### **3. Архитектура современных и перспективных АСДУ**

Современные АСДУ имеют многоуровневую архитектуру [12]. Нижний уровень образуют периферийные устройства и инженерное оборудование, формирующие первичные данные. Второй уровень – контроллеры, принимающие и обрабатывающие информацию, и сеть передачи данных. Верхний уровень – это программное обеспечение, предоставляющее средства визуализации, архивации, публикации поступающих данных, а также средства управления бизнес-процессами. На рабочие места диспетчеров (АРМ) поступает структурированная консолидированная информация в нужном формате. Аналитический модуль постоянно отслеживает рабочие параметры систем на предмет отклонения от нормы и способен автоматически запускать процедуры согласно заложенным инструкциям, например, подать сигнал тревоги или запустить резервное оборудование.

АСДУ последних поколений представляют собой целостные платформы для управления всеми инженерными подсистемами и создаются как многоуровневые автоматические системы, обеспечивающие контроль состояния и управление технологическим оборудованием с выводом данных на экраны АРМ операторов, ведением непрерывного мониторинга инженерных систем с регистрацией основных параметров и обеспечивают контроль и управление инженерным комплексом из единого диспетчерского центра.

Любая современная АСДУ способна централизованно фиксировать события в базе данных и оповещать диспетчера о возникновении проблемы и необходимости ее разрешения. Далее система может определить уровень серьезности происшествия и присваивает событию определенный приоритет. Приоритет необходим, чтобы повысить эффективность реакции персонала на происшествие. Система выводит сообщения о выходе отслеживаемых параметров за установленные ранее пределы, а также сообщения о критическом времени наработки эксплуатируемого инженерного оборудования. Информация представляется для администраторов и диспетчеров в легко читаемом виде.

Алгоритмы и регламенты ответных действий на произошедшее событие программируются в АСДУ, и от правильности настройки подобных регламентов напрямую зависит эксплуатационная готовность. Для разграничения ответственности за обслуживание разных систем АСДУ имеет возможность управлять полномочиями диспетчеров. Автоматизированная система предоставляет функции разграничения доступа различных групп диспетчеров с привязкой к определенным задачам или контролируемым системам. Аналитическая часть АСДУ предоставляет необходимый инструмент для определения причин простоев и планирования уровня избыточности инженерных систем, а также содержит экспертные подсистемы для принятия решений.

Наиболее известными перспективными решениями АСДУ с наибольшим количеством внедрений на крупных объектах энергетических компаний в России и мире являются: Network Manager (ABB), Spectrum PowerCC и Energy Management Suite (Siemens), GENe EMS (SNC-Lavalin), PACiS (Areva), СК-2007С и СК-2003 (Монитор Электрик).

### Заключение

Особенности энергосистем и электросетей России зачастую не позволяют механически копировать зарубежный опыт создания и развития информационных технологий. В первую очередь, это связано с недостаточностью объема оперативной информации, автоматически доставляемой на диспетчерские пункты. В условиях недостатка оперативной информации многие элементы «традиционных» информационных технологий не в состоянии эффективно функционировать. Поэтому для развития информационных технологий в отечественной энергетике актуально развитие новых технологий и адаптация зарубежных технологий к российским условиям.

Неполнота информации о состоянии ЕНЭС (и, прежде всего, неполнота данных по топологии электрической сети) затрудняет использование разработанных технологических модулей в онлайн-режиме. Поэтому целесообразно дополнение «традиционных» технологических программных комплексов мониторинга новыми разработками на основе технологии экспертных систем. Важная задача проработки ремонтных заявок также может быть эффективно решена с использованием технологии экспертных систем.

Все известные АСДУ имеют практически однотипный набор базовых функций и возможностей, отличия определяются типом используемых баз данных, серверных и клиентских операционных систем, возможностями поддержки оборудования предыдущих поколений, а также организацией средств коммуникационного взаимодействия.

При реализации технологии Smart Grid необходимо учитывать особенности структуры линий электропередач и рынка электроэнергии в целом [13].

Таким образом, для внедрения новых технологий мониторинга и управления достаточно улучшать существующую инфраструктуру, устанавливая "интеллектуальные" счетчики. В некоторых отдельных случаях требуется вначале сменить оборудование для электропередачи, которое морально и физически устарело, и только затем начинать комплексное внедрение новых технологий.

### Литература

1. Автоматизация диспетчерского управления в электроэнергетике / В.А. Баринов и др. М.: Изд-во МЭИ, 2000.
2. Егоров В., Кужеков С. Интеллектуальные технологии в распределительном электросетевом комплексе // ЭнергоРынок, №6, 2010.
3. Лондер М.И., Тумаков А.В. Единое информационное пространство как основа создания интегрированной системы управления электрическими сетями России // Естественные и технические науки. 2010. № 4.
4. Любарский Ю.Я., Моржин Ю.И. Отечественные оперативно-информационные комплексы АСДУ энергосистем // Электрические станции. 2001. № 12.
5. Орнов В.Г., Моржин Ю.И. Развитие оперативно-информационных комплексов автоматизированных систем диспетчерского управления в России // Электрические станции. 2005. № 1.
6. Автоматизация диспетчерского управления ЕЭС России / Ю.И. Моржин и др. // Энергетик. 2001. № 2.
7. Моржин Ю.И., Цветков Е.В. Совершенствование систем автоматизированного диспетчерского управления в нормальных и аварийных режимах. Новые технологии в энергетике // Сборник РАО «ЕЭС России». 2002.
8. Моржин Ю.И., Лондер М.И., Чепкасов А.П. Концепция создания единой информационной среды интегрированной системы управления электрическими сетями России // Вестник ВНИИЭ, 2004.
9. Колосок И.Н. Современные направления в области развития методов оценивания состояния ЭЭС и их реализации при оценивании состояния по контрольным уравнениям // Управление энергетическими системами – новые технологии и рынок : сб. Сыктывкар, 2004.
10. Любарский Ю.Я., Моржин Ю.И. Концепция «интеллектуальных» оперативно-информационных комплексов для АСДУ энергосистем // Инновации в энергетических технологиях : сб. – М.: Агро - Принт, 2002.
11. Руденко Ю.Н., Семенов В.А. Автоматизация диспетчерского управления в энергетике. – М.: Изд-во МЭИ, 2000.
12. Многоуровневая АСДУ в энергетике - современные тенденции в организации и разработке / А.Л. Вулис и др. // Электрические станции. 1994. № 9.
13. URL: [http://www.cleandex.ru/articles/2010/04/13/Smart\\_Grid\\_market](http://www.cleandex.ru/articles/2010/04/13/Smart_Grid_market)

*Сахабетдинов Ильдар Умарович*, научный сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, тел. (495) 3347831, e-mail: [yonavi@ipu.ru](mailto:yonavi@ipu.ru)

*Sakhabetdinov Ildar Umarovich*, scientific researcher, V.A. Trapeznikov Institute of control sciences of RAS



УДК 004.4'244

© *Е.С. Фереферов, Ю.А. Новицкий, Г.М. Ружников, А.Е. Хмельнов*

## **ТЕХНОЛОГИЯ ИНТЕГРАЦИИ ГЕОИНФОРМАЦИОННЫХ ФУНКЦИЙ В ИНФОРМАЦИОННЫЕ СИСТЕМЫ<sup>1</sup>**

Статья посвящена проблемам расширения функциональных возможностей существующих информационных систем (ИС) за счёт встраивания геоинформационных функций. Предложен ряд инструментальных средств, автоматизирующих данный процесс.

**Ключевые слова:** информационные системы, геоинформационные системы, инструментальные средства.

*E.S. Fereferov, Yu.A. Novitsky, G.M. Ruzhnikov, A.E. Khmelnov*

## **TECHNOLOGY OF INTEGRATION OF GIS FUNCTIONS TO INFORMATION SYSTEMS**

The article is devoted to the problems of extending the functionality of existing information systems (IS) by embedding GIS functions. The authors propose tool systems to automate that process.

**Keywords:** information systems, geographic information systems, tool systems.

### **Введение**

В информационных системах, ориентированных на управление территориальным развитием, как правило, используется информация, имеющая пространственную привязку (например, почтовые адреса, кадастровые номера). При этом в таких системах часто отсутствует геоинформационная составляющая. Применение ГИС технологий повышает наглядность представления информации, позволяет проводить пространственный анализ и как следствие позволяет повысить качество и обоснованность управленческих решений. Поэтому расширение функциональных возможностей существующих ИС за счёт интеграции в них ГИС-функций является актуальной задачей.

Для решения данной задачи существует несколько подходов: разработка ИС средствами ГИС (например, «Приложение БД» в ГИС «Панорама» [1]) или реализация функций ГИС в существующих АИС. В первом случае приходится реализовывать все функции существующей информационной системы средствами ГИС, по сути, создавая систему заново. При этом программисту может потребоваться освоение встроенных языков ГИС (например, Avenue для ArcView [2], MapBasic для MapInfo [3]). Второй подход требует от разработчика знаний в области геоинформационных технологий, поскольку необходима реализация функций для работы с пространственными объектами, а также обязательное наличие исходных текстов существующей ИС.

Из вышесказанного можно сделать вывод, что интеграция геоинформационных функций в существующие ИС – задача трудоемкая. Поэтому актуальна разработка инструментальных средств, позволяющих автоматизировать интеграцию функций ГИС в существующие системы.

### **Картографический модуль**

Для работы с цифровыми картами из внешних приложений авторами создан картографический модуль, реализованный в виде СОМ-сервера, к которому могут подключаться ИС-клиенты (возможно несколько сразу). Реализация модуля для работы с ГИС в виде СОМ-сервера дает ряд полезных свойств. Во-первых, при взаимодействии с несколькими приложениями-клиентами, картографический модуль является интегрирующей составляющей, позволяющей объединять информацию из различных информационных систем на уровне ГИС. Во-вторых, существует возможность средствами ОС Windows настроить запуск картографического модуля от имени пользователя, обладающего правами доступа к файлам карты, что позволяет исключить несанкционированное копирование файла электронной карты. В режиме «по умолчанию» СОМ-приложения запускаются под учетной записью текущего (запускающего) пользователя, но есть

---

<sup>1</sup> Работа выполнена при частичной финансовой поддержке РФФИ проекты 11-07-00426-а, 11-07-92204-Монг-а, междисциплинарного проекта № 131 СО РАН, ОНИТ РАН (проект № 4.1)

возможность при помощи утилиты dcomcnfg.exe настроить запуск COM-сервера под учетной записью указанного пользователя. Данная возможность позволяет наладить работу с файлами карт, доступ к которым текущему пользователю запрещен.

Методы интерфейса приложения созданного картографического модуля позволяют вызывать окно для работы с цифровыми картами (рис.1) в формате ГИС «Панорама» [1], и обеспечивают выполнение ряда ГИС-функций. Применение данного модуля при разработке или модернизации существующих ИС позволяет избежать необходимости написания функций для работы с цифровой картой. В картографическом модуле уже реализованы следующие ГИС-функции:

- масштабирование,
- перетаскивание,
- получение информации по объекту карты (семантика и метрика объекта),
- создание новых объектов,
- создание пользовательских слоев,
- поиск объектов на цифровой карте по адресу (функция «Адресный план»).

Для взаимодействия с ИС-клиентами библиотека типов созданного COM-сервера содержит описание двух интерфейсов: IApplication и IClient. IApplication содержит функции для работы с цифровыми картами, реализованные в картографическом модуле, а IClient – функции, вызываемые картографическим модулем для передачи или получения информации из программы-клиента, которые должны быть реализованы на клиенте соответственно. Рассмотрим данные интерфейсы более подробно.

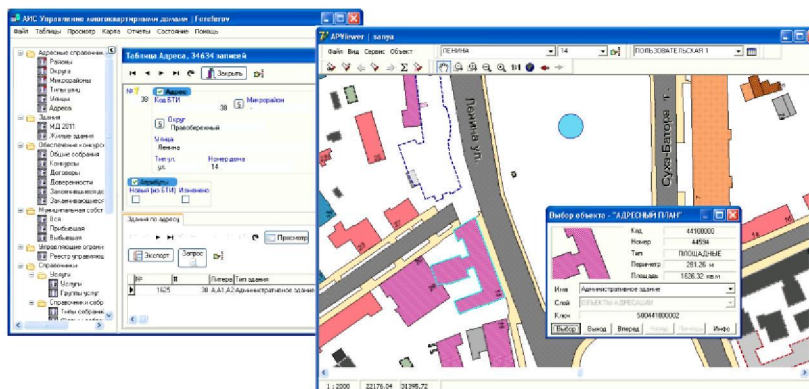


Рис. 1. Взаимодействие ИС с картографическим модулем

Методы интерфейса IApplication:

**function** SetClientInt(const ClientInt: IClient): HRESULT – функция передачи клиентского интерфейса.

**function** AddClientButton(Tg: Integer; const Hnt: WideString; SN: Integer): HRESULT – функция позволяет создать в окне картографического модуля кнопку с номером (Tg), подписью (Hnt), при нажатии которой в приложение-клиент будет передано значение семантики с номером SN выбранного объекта. Созданным клиентским кнопкам может быть задан внешний вид.

**function** ShowObjByAddr(const S: WideString; const H: WideString): HRESULT – функция позволяет искать объекты цифровой карты, у которых семантика «Улица»=S, «Дом»=H. При этом объекты удовлетворяющие запросу выделяются цветом и перемещаются в центр окна (происходит фокусировка).

**function** ShowObjsByAddrs(const Addrs: WideString; L: Integer): HRESULT – Функция поиска объектов на карте по списку адресов вида «Улица1, Дом1, Улица2, Дом2, ..., УлицаL, ДомL».

**function** ShowObjByID(ID: Integer): HRESULT – Функция поиска объектов на цифровой карте по уникальному ключу ID объекта.

**function** ShowObjsBySemNum(SN: Integer; const SV: WideString): HRESULT – функция позволяет искать объекты цифровой карты, у которых семантика с номером SN имеет значение SV.

**function** AddObjsToList(IDL: Integer; IDOBJ: Integer; SelAppend: Integer; Color: Integer): HRESULT – функция добавляет в список выборки объект с номером пользовательской карты IDL, уникальным ключом IDOBJ и цветом выделения Color. При этом, если параметр SelAppend=1, то

список выборки будет создан заново иначе объекты будут добавляться в существующий список. Для отображения списка объектов необходимо вызвать **function ShowSelList: HRESULT**.

Методы интерфейса **IClient**:

**procedure GetObjInfoPress(Tg: Integer; const VL: WideString)** – процедура для обработки получаемых значений семантики VL выбранного объекта карты при нажатии пользовательской кнопки с номером Tg в картографическом модуле.

**function GetFieldValByNameClnt(const FldName: WideString): WideString** – функция для получения значения из поля FldName таблицы. Данная функция может вызываться картографическим модулем при построении отчетов, содержащих связанную с объектом карты информацию из БД приложения-клиента.

**procedure ShowMapObjDBInfo(hL: Integer; hObj: Integer)** процедура для поиска и отображения в приложении-клиенте информации связанной с выбранным объектом карты.

Описанные процедуры и функции в большинстве случаев достаточны для решения задачи взаимодействия ИС-клиентов с картографическим модулем.

### Модуль формирования отчетов по карте

Для поддержания возможности создания отчетов, содержащих информацию об объектах цифровых карт, авторами был разработан генератор отчетов, реализованный в виде динамически подключаемой библиотеки (dll), взаимодействующей с картографическим модулем. В основе данного построителя лежит технология создания отчетов с использованием метаописаний структур БД и шаблонов в MS Word, содержащих метки форматирования данных[4], впервые апробированная при создании генератора отчетов для системы ГеоАРМ[5].

Изначально генератор был разработан для построения отчетов с использованием метаописаний структур баз данных. При вызове генератора из метаописания ему передавалась информация о соединении с БД, файле-шаблоне и таблице или представлении (таблице со значениями полученными по ссылкам), для записей из которых нужно построить отчет. В метаописании таблиц (представлений) могла содержаться информация о связях типа «Мастер-Детали», что позволяет наладить вывод данных из таблиц-деталей.

При создании отчетов по объектам цифровой карты, семантика объекта воспринимается генератором отчетов как поля записи таблицы, а метрика как детали. При этом именами полей выступают имена соответствующих семантик.

В процессе подготовки отчета происходит нахождение меток в тексте шаблона, после чего на основании параметров метки происходит вставка данных. Шаблоны отчетов содержат метки следующего вида:

<#**[Вид данных]** N=**Имя (Формат)**>, где

**[Вид данных]** = {FV – значение семантики; IMG – изображение фрагмента карты; DETAIL – значения из метрики};

**N=Имя** – имя соответствующей семантики, из которой будут браться данные;

**(Формат)** – необязательный набор параметров, для каждого вида данных свой. Например, для склонения подставляемых значений можно указать падеж (для реализации данной функции была применена библиотеки padeg.dll[6]). Имя может иметь вид: N = «^» «Имя таблицы» «.»

«Имя поля» – тогда данные будут подставляться из соответствующего поля и таблицы приложения клиента.

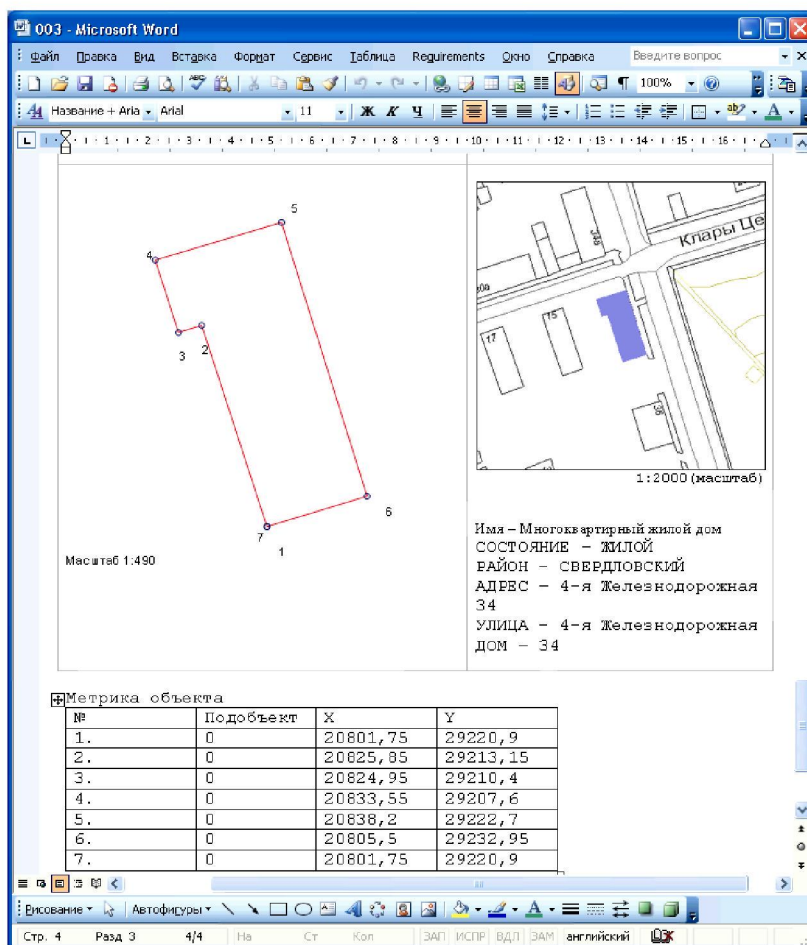


Рис. 2. Отчёт с объектами цифровой карты

Кроме того в отчёте может быть опубликован объект цифровой карты с помощью примитивов графики MS Word (рис. 2). Для этого в шаблоне документа должны присутствовать объекты MS Word: полотно с именем «Границы», задающее область рисования, автофигура с именем «Центр», задающая центр рисования объекта, а также автофигуры, отвечающие за служебные подписи.

### Заключение

Описанные выше инструментальные средства позволяют интегрировать функции для работы с векторными картами в формате «Панорама» в приложения, где такая работа не была ранее поддержана. Реализация картографического модуля в виде СОМ-сервера позволяет регламентировать работу с электронной топоосновой на уровне политик безопасности ОС Windows, при этом исключив несанкционированное копирование файлов карт. Модуль публикации позволяет создавать отчётные формы, содержащие информацию об объектах цифровых карт. Разработанные инструментальные системы были внедрены в МУП Бюро технической инвентаризации г. Иркутска.

### Литература

1. ГИС Панорама. URL: <http://www.gisinfo.ru>
2. GIS and Mapping Software. URL: <http://esri.com>
3. MapInfo Corporation. URL: <http://mapinfo.com>
4. Фереферов Е. С., Хмельнов А. Е. Технология создания отчетов с использованием метаописаний структур БД и шаблонов, содержащих метки форматирования данных. Ляпуновские чтения: материалы конф. Иркутск, 2008. С. 53-54.
5. Фереферов Е.С., Бычков И.В., Хмельнов А.Е. Метаописание баз данных как основа интеграции информационно-справочных систем и ГИС // «Вычислительные технологии». Т.12, № 5. С. 41-51.

6. Склонение фамилий, имен и отчеств по падежам. Библиотека функций. Виртуальный клуб программистов «Королевство Delphi». URL: <http://www.delphikingdom.com/asp/viewitem.asp?catalogid=412>

*Фереферов Евгений Сергеевич*, научный сотрудник ИДСТУ СО РАН, 664033, г. Иркутск, ул. Лермонтова, 134, тел. (395-2) 453071, e-mail: [fereferov@icc.ru](mailto:fereferov@icc.ru)

*Новицкий Юрий Александрович*, программист ИДСТУ СО РАН, тел. (395-2) 453071, e-mail: [novitskiy@icc.ru](mailto:novitskiy@icc.ru)

*Ружников Геннадий Михайлович*, канд. техн. наук, ИДСТУ СО РАН, тел. (395-2) 453006, e-mail: [ruginov@icc.ru](mailto:ruginov@icc.ru)

*Хмельнов Алексей Евгеньевич*, канд. техн. наук, зав. лаб. ИДСТУ СО РАН, тел. (395-2) 453071, e-mail: [hmelnov@icc.ru](mailto:hmelnov@icc.ru)

*Fereferov Evgeniy Sergeevich*, researcher of Institute for System Dynamics and Control Theory of SB RAS, e-mail: [fereferov@icc.ru](mailto:fereferov@icc.ru)

*Novitskiy Yuriy Alexandrovich*, programmer of Institute for System Dynamics and Control Theory of SB RAS

*Ruzhnikov Gennadiy Mikhailovich*, candidate of technical science, Institute for System Dynamics and Control Theory of SB RAS

*Khmelnov Alexey Evgenievich*, candidate of technical science, Chief of laboratory of Institute for System Dynamics and Control Theory of SB RAS



УДК 621.391

© М.В. Харинов

## УСТОЙЧИВАЯ СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЯ<sup>1</sup>

Статья посвящена так называемой «проблеме сегментации», возникающей в области автоматического распознавания цифровых изображений на стадии выделения визуально воспринимаемых объектов. Для достоверного выделения и распознавания объектов в нашем подходе известный гистограммный метод Оцу (1979) объединяется с моделью Мамфорда – Шаха сегментации изображения по локальным признакам (1985), которая опирается на понятие оптимального приближения изображения ступенчатой (кусочно-постоянной) функцией. В статье, на стандартном примере реального изображения, демонстрируются оптимальные приближения, полученные перебором пороговых значений по участкам яркостной гистограммы. Решается задача исключения перебора вариантов за счет учета пространственного распределения пикселей в изображении. Для получения решения определяется условие устойчивой сегментации, которое тестируется на стандартном примере.

**Ключевые слова:** сегментация, среднее квадратичное отклонение, минимизация, оптимальные приближения, перекрывающиеся разбиения, слияние сегментов, условие устойчивости.

*M.V. Kharinov*

## STABLE IMAGE SEGMENTATION

The article is devoted to so-called Segmentation Problem that occurs in domain of automatic recognition of digital images when detecting of visually perceived objects. For proved object detection and recognition the famous histogram Otsu's method (1979) is combined with Mumford–Shah model (1985) of local (bottom up) segmentation basing on the optimal piecewise constant image approximations. The article presents the optimal approximations for the standard example of the real image obtained by exhaustive search of histogram thresholds in limited intensity ranges. To avoid an exhaustive search the spatial distribution of pixels in an image is taken into account. At that the stability condition for the segmented images is derived and tested on the standard image.

**Keywords:** segmentation, standard deviation, minimization, optimal approximations, overlapping partitions, segment merging, stability condition.

### Введение

В области компьютерного распознавания изображений ключевой проблемой является формальное определение понятия «объектов», которые должен выделить компьютер прежде, чем приступить к выделению признаков и идентификации объектов развитыми средствами распознавания образов [1, 2]. Под выделением объектов понимается указание занимаемых ими участков изображения, которые обычно задают перечислением связанных сегментов, вычисляемых в результате сегментации.

В современной литературе проблему сегментации часто относят к слабо формализованным, что справедливо, если сегментация определяется не как результат, а как процесс разбиения пикселей изображения на множества по некоторому алгоритму. В этом случае, результирующее разбиение зависит от особенностей входных данных, что приводит к разделению изображений на ряд предметных областей (телевизионных, аэрокосмических, медицинских и пр.), для которых разрабатываются конкретные алгоритмы сегментации. В более точной постановке проблема сегментации состоит в определении сегментированного изображения независимо от конкретного алгоритма или технологии его получения. Другими словами, определить сегментированное изображение – значит указать, чем оно отличается от изображения, просто разделенного на части по тому или иному алгоритму. При этом достоверность сегментации можно проверить формально без использования понятия зрительно воспринимаемых объектов, а выделение объектов трактовать как неформальную интерпретацию сегментации.

Самосогласованная постановка задачи сегментации обеспечивается определением сегментированного изображения как оптимального приближения, минимально отличающимся от изображения по значению некоторого функционала.

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ, проект №11-07-00685-а.

### 1. Постановка задачи

В отличие от большинства работ по оптимальной сегментации, мы не ограничиваемся поиском единственного оптимального приближения изображения, а решаем задачу получения оптимальных приближений для каждого числа множеств, на которые разбивается множество пикселей изображения, причем результирующая последовательность разбиений пикселей изображения на множества не предполагается иерархической.

В методе Оцу [3] строится оптимальное разбиение пикселей изображения из  $N$  пикселей на два множества (класса), отвечающее минимальному значению внутриклассовой дисперсии. В тривиальном обобщении постановки задачи на случай последовательного разбиения пикселей по яркости на  $n = 1, 2, \dots, N$  градаций, минимизируемый функционал  $\sigma^2$  записывается как взвешенная сумма дисперсий  $\sigma_i^2$  по множествам  $i$ :

$$\sigma^2 = \sum_{i=1}^n \omega_i \sigma_i^2, \quad (1)$$

где  $\omega_i$  – вероятность, сопоставляемая множеству  $i$ .

Для изображения из  $N$  пикселей величина  $N\sigma^2$  есть сумма квадратов отклонений яркостей пикселей от средних значений, вычисленных по каждому множеству пикселей  $i$ . В модели Мамфорда–Шаха [4–9], в которой рассматриваются множества пикселей связанных сегментов изображения, эта величина является основной компонентой «энергетического» функционала [6]:

$$N\sigma^2 + \lambda L, \quad (2)$$

где  $L$  – суммарная длина границ между сегментами, а  $\lambda$  – так называемый, «регуляризационный» параметр.

Величина  $\sigma$  в (2) описывает среднеквадратичное отклонение изображения от своего кусочно-постоянного приближения с усредненными по сегментам значениями яркости пикселей.

В современных практических реализациях модели Мамфорда – Шаха [8–9] регуляризационный параметр  $\lambda$  рассчитывается автоматически, но оказывается зависим от изображения, что нарушает идею минимизации функционала (2). В [7] указанный параметр полагается равным 0, что не противоречит (2) и сохраняет нетривиальный смысл минимизации  $\sigma$ , если она выполняется для каждого числа сегментов.

По нашему опыту, учет границ сегментов при нулевом параметре  $\lambda$  существенно не влияет на практические результаты, а для их улучшения важнее учесть, что последовательность оптимальных приближений изображения, вообще говоря, не является иерархической. Поэтому операция слияния сегментов, которой ограничиваются в модели Мамфорда – Шаха [4–9], оказывается недостаточной.

Для получения кусочно-постоянных приближений, минимально отличающихся от изображения по среднеквадратичному отклонению  $\sigma$  или квадратичной ошибке  $E \equiv N\sigma^2$ :

$$\sigma = \min \Leftrightarrow E = \min, \quad (3)$$

вводится дополнительная операция коррекции множеств пикселей изображения. При этом связность пикселей в искомым множествах, в общем случае, не считается обязательной, как в методе Оцу.

### 2. Критерии сегментации и условие устойчивости

В рамках модели Мамфорда – Шаха сегментация выполняется посредством итеративного слияния рассматриваемых множеств пикселей. Слияние множеств 1 и 2 выполняется при минимальной величине приращения квадратичной ошибки  $\Delta E_{merge}$ , которое в версии [7] модели записывается в виде:

$$\Delta E_{merge} \equiv \frac{(I_1 - I_2)^2}{\frac{1}{n_1} + \frac{1}{n_2}} = \min, \quad (4)$$

где  $I_1, I_2$  – средние значения пикселей, а  $n_1, n_2$  – число пикселей в множествах 1 и 2.

Коррекция множеств состоит в реклассификации пикселей одинаковой яркости  $I$ , исключаемых из множества 1 и включаемых в число пикселей множества 2. Критерием коррекции является

отрицательная величина сопутствующего приращения квадратичной ошибки  $\Delta E_{correct}$ :

$$\Delta E_{correct} \equiv \frac{(I_1 - I_2)^2}{\frac{1}{k} + \frac{1}{n_2}} - \frac{(I - I_1)^2}{\frac{1}{k} - \frac{1}{n_1}} < 0, \quad (5)$$

где  $k$  – число реклассифицируемых пикселей.

Элементарным преобразованием из (5) получаем условие коррекции множеств, учитывающее только знак  $\Delta SE_{correct}$ :

$$|I - I_1| > \alpha \cdot |I - I_2|, \quad (6)$$

где коэффициент  $\alpha < 1$  описывает соотношение количеств пикселей в рассматриваемых множествах и определяется в виде:

$$\alpha \equiv \sqrt{\frac{n_2(n_1 - k)}{n_1(n_2 + k)}}. \quad (7)$$

Из (6), (7) для реклассификации  $k$  пикселей одинаковой яркости легко установить, что если среднеквадратичное отклонение  $\sigma$  снижается при реклассификации одного пикселя, то оно снижается и при реклассификации остальных пикселей той же яркости, поскольку с ростом  $k$  коэффициент  $\alpha$  уменьшается.

Применяя логическое отрицание к выражениям (5), (6), получаем условие устойчивости:

$$\Delta E_{correct} \geq 0 \Leftrightarrow |I - I_1| \leq \alpha \cdot |I - I_2|. \quad (8)$$

Разбиение пикселей изображения на множества считается устойчивым, если для всех рассматриваемых пар множеств реклассификация пикселей одинаковой яркости из одного множества в другое не приводит к повышению среднеквадратичного отклонения  $\sigma$ .

Очевидно, оптимальное разбиение изображения, минимально отличающееся от изображения по  $\sigma$ , является устойчивым, но устойчивое разбиение не обязательно оптимально. Ограничения на допустимые пары множеств 1 и 2, например, обменивающихся пикселями при сохранении связности, или рассмотрение одновременной реклассификации двух и более подмножеств пикселей приводит к различным вариантам программной реализации условия устойчивости (8).

### 3. Экспериментальные результаты

Рис. 1 иллюстрирует оптимальные приближения, вычисленные для стандартного изображения «Лена».

Исходное изображение показано в правом нижнем углу. Остальные 7 изображений являются его кусочно-постоянными приближениями в последовательно возрастающем числе градаций яркости, размещенными в порядке возрастания числа градаций от 1 до 7 слева направо и сверху вниз.

Всего приближений 216, так как в изображении «Лена» встречаются пиксели 216 градаций яркости. Для приближений в  $n = 1 - 64$  яркостных градациях значения среднеквадратичного отклонения  $\sigma$  перечислены в столбцах таблицы 1 – по 16 штук в каждом столбце.





Рис. 1. Оптимальные приближения изображения

Таблица 1

Оценки  $\sigma$  для оптимальных приближений изображения «Лена» в  $n$  градациях яркости

$n$	$\sigma$	$n$	$\sigma$	$n$	$\sigma$	$n$	$\sigma$
1	55.88322	17	3.70326	33	1.72854	49	1.21323
2	30.64564	18	3.50441	34	1.68547	50	1.18618
3	21.21739	19	3.32383	35	1.64729	51	1.16091
4	14.96450	20	3.15658	36	1.61313	52	1.13576
5	11.69762	21	3.01260	37	1.57894	53	1.11091
6	10.03975	22	2.87305	38	1.54589	54	1.08554
7	8.46072	23	2.73844	39	1.51218	55	1.06012
8	7.51121	24	2.60572	40	1.47850	56	1.03435
9	6.81359	25	2.49076	41	1.44586	57	1.00904
10	6.14397	26	2.37391	42	1.41400	58	0.98520
11	5.57864	27	2.25584	43	1.38327	59	0.96397
12	5.11403	28	2.15389	44	1.35340	60	0.94304
13	4.75689	29	2.05655	45	1.32287	61	0.92172
14	4.42306	30	1.95565	46	1.29542	62	0.90113
15	4.17825	31	1.87573	47	1.26794	63	0.88054
16	3.92460	32	1.79485	48	1.23991	64	0.86341

Для всей последовательности приближений зависимость среднеквадратичного отклонения  $\sigma$  от числа градаций  $n$  в логарифмическом масштабе по горизонтальной оси иллюстрируется графиком на рис. 2.

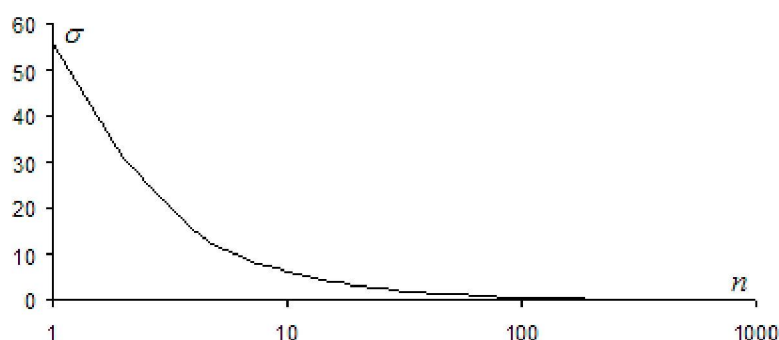


Рис. 2. График зависимости оценок  $\sigma$  от числа градаций  $n$  для оптимальных приближений изображения «Лена»

Характерно, что последовательность оптимальных приближений изображения не является иерархической. При этом сегменты из различных разбиений перекрываются между собой (рис. 1). Строго говоря, обсуждаемые «оптимальные» приближения изображения являются правдоподобными оценками точных приближений, минимально отличающимися от изображения по среднеквадратичному отклонению  $\sigma$  на множестве доступных алгоритмов. Они вычислены независимо от линейных преобразований изображения по яркости и преобразования изображения из позитива в негатив посредством ограниченного перебора вариантов.

Минимизация  $\sigma$  достигалась итеративным повторением обработки всей последовательности разбиений до получения неизменного результата. Каждое разбиение преобразовывалась из условия оптимального разбиения на градации подмножеств пикселей по всем участкам гистограммы изображения, отвечающим не более чем  $k$  последовательным градациям яркости ( $k = 2, 3, \dots$ ). Оптимальные пороговые значения яркости находились перебором вариантов. Посредством слияния/разбиения текущих диапазонов пикселей задавались преобразования разбиений одного в другое с изменением числа градаций на единицу. Изменение текущего значения среднеквадратичного отклонения  $\sigma$  и устойчивость получаемых разбиений контролировались по формулам (4)-(8).

Для исключения перебора вариантов за счет учета пространственного распределения в изображении пикселей одинаковой яркости мы применили модель Мамфорда – Шаха, которую программно реализовали в приложении к любым множествам пикселей, рассматриваемым вместо традиционных связанных сегментов изображения<sup>2</sup>. При этом геометрические свойства рассматриваемых множеств выражаются отношением смежности, которое устанавливается для множеств так же, как и для связанных сегментов изображения.

Вычисление последовательных разбиений с учетом геометрического распределения пикселей проводится итеративно и сводится к слиянию множеств пикселей согласно (4) по очереди с их коррекцией, выполняемой до получения устойчивого разбиения изображения (8). В процессе коррекции смежные множества пикселей обмениваются подмножествами пикселей одинаковой яркости из условия падения среднеквадратичного отклонения (5)-(7).

Эксперимент показал что, если в качестве начального разбиения на множества берется разбиение пикселей изображения по градациям исходной яркости, то получаемые разбиения при достаточной редукции числа множеств пикселей воспроизводят результаты вычислений, полученные перебором вариантов по методу Оцу. Так, в случае стандартного изображения (рис. 1), слияние смежных множеств пикселей по очереди с коррекцией получаемых разбиений обеспечивает воспроизведение приближений в 1–15 градациях яркости (табл.). График зависимости среднеквадратичного отклонения от числа множеств пикселей, вычисленных в модели Мамфорда–Шаха, сливается с графиком среднеквадратичного отклонения в зависимости от числа градаций яркости рис. 2. По визуальному восприятию сравниваемые приближения практически совпадают между собой,

<sup>2</sup> Случай связанных сегментов рассмотрен в [10]. Структура данных и детали вычислений описываются в [11–12].

так как при более, чем 15 множествах пикселей (градациях яркости) заметным образом не отличаются от исходного изображения.

### **Заключение**

Таким образом, в статье аналитически и экспериментально обоснован подход к сегментации цифровых изображений посредством оптимальных кусочно-постоянных приближений, характеризующихся минимальными значениями среднеквадратичного отклонения приближения от изображения. В качестве основы вычислений использованы метод Оцу и модель Мамфорда–Шаха, которые обобщены и развиты применительно к условиям рассматриваемой задачи.

При обобщении метода Оцу разработан и программно реализован гистограммный способ вычисления оптимальных приближений изображения посредством ограниченного перебора пороговых значений для каждого числа градаций яркости. В качестве альтернативного способа вычисления оптимальных приближений предложено использовать модель Мамфорда–Шаха с редуцированным функционалом (3), операцией коррекции (5)–(7), дополняющей слияние множеств пикселей, и обобщенной трактовкой самих множеств, к которым не предъявляется условие связности. В качестве формального признака оптимального разбиения установлено свойство устойчивости (8) сегментации относительно предусмотренных вариантов реклассификации пикселей изображения.

Полученные результаты по формализации понятия сегментации планируется применить для унификации программного обеспечения выделения объектов на цифровых изображениях произвольного содержания.

### **Литература**

1. Ту Дж., Гонсалес Р. Принципы распознавания образов. – М.: Мир, 1978. – 411 с.
2. Журавлев Ю.И. Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики, – М.: Наука, 1978. Вып. 33. С. 5-68.
3. Otsu N. A Threshold Selection Method from Gray-Level Histograms // IEEE Transactions on systems, MAN, and CYBERNETICS, January 1979. Vol. SMC-9, № 1, P. 62-66.
4. Mumford D., Shah J. Boundary detection by minimizing functionals, I // Proc. IEEE Comput. Vision Patt. Recogn. Conf., San Francisco. 1985. P. 22-26.
5. Mumford D., Shah J. Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems // Communications on Pure and Applied Mathematics. 1989. Vol. XLII, № 4. P. 577-685.
6. Koepfler G., Lopez C., Morel J. A Multiscale Algorithm for Image Segmentation by Variational Method // SIAM Journal on Numerical Analysis. 1994. Vol. 31, № 1. P. 282-299.
7. Бугаев А.С., Хельвас А.В. Поисковые исследования и разработка методов и средств анализа и автоматического распознавания потоковой информации в глобальных информационных системах. Шифр «Лацкан» // Отчет по НИР. М.: МФТИ, 2001. Т. 1. 140 с.
8. Redding N.J., Crisp D.J., Tang D.H., Newsam G.N. An efficient algorithm for Mumford–Shah segmentation and its application to SAR imagery // Proc. Conf. Digital Image Computing Techniques and Applications (DICTA '99). 1999. P. 35-41.
9. Robinson B.J., Redding N.J., Crisp D.J. Implementation of a fast algorithm for segmenting SAR imagery // Scientific and Technical Report, Australia: Defense Science and Technology Organization, Australia. 01 January 2002. – 42 p.
10. Kharinov M. V. Merge-and-correct segmentation method // Proc. of the 8th Open German-Russian Workshop on Pattern Recognition and Image Understanding (OGRW-8-2011). Nizhny Novgorod: Nizhny Novgorod Lobachevsky State University, November 21-26 2011. P. 123-126.
11. Kharinov M.V. Adaptive Dichotomous Image Segmentation Toolkit // Pattern Recognition and Image Analysis: Advances in Mathematical Theory and Applications. 2012. Vol. 22, № 1. P. 228-235.
12. Харинов М.В. Запоминание и адаптивная обработка информации цифровых изображений. – СПб.: Изд-во С.Петербург. ун-та, 2006. – 138 с.

*Харинов Михаил Вячеславович*, канд. техн. наук, старший научный сотрудник лаборатории Прикладной информатики Санкт-Петербургского института информатики и автоматизации РАН, тел. (812) 3281919, e-mail: [khar@iias.spb.su](mailto:khar@iias.spb.su)

*Kharinov Mikhail Vyacheslavovich*, candidate of technical science, Senior researcher of Laboratory of Applied Informatics of the St. Petersburg Institute for Informatics and Automation of RAS

УДК 519.683+519.688+519.71+519.1

© М.Ю. Чернышов, Н.В. Абасов, Е.Н. Осипчук

## ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНОЛОГИИ ДЛЯ АНАЛИЗА И РЕСТРУКТУРИРОВАНИЯ ТЕКСТОВ ПРОГРАММ НА ПРИНЦИПАХ ЛОГИКО-СМЫСЛОВОГО АНАЛИЗА И СИНТЕЗА

Излагаются принципы технологии вычислительного моделирования, которая рассчитана на эффективный анализ неизвестных программ и реструктурирование их текстов. На ее основе разработан программный комплекс, предназначенный для содержательного анализа кода без выполнения программы. Технология анализа предполагает автоматическое построение графов связей между функционально-семантическими кластерами в текстах программ. Глубина такого анализа может варьировать от определения поведения (функций) отдельных операторов до анализа всего исходного кода, а его результаты могут быть использованы для оперативного выявления ошибок в программе или проверки ее соответствия спецификации. Технология реструктурирования текстов программ допускает перестройку текстов по заданному алгоритму и пополнение графов связей дополнительными кластерами.

**Ключевые слова:** технологии вычислительного моделирования; содержательный анализ текстов программ; реструктурирование текстов программ; статический анализ кода; графы связей.

*M.Yu. Chernyshov, N.V. Abasov, E.N. Osipchuk*

## FOUNDATIONS OF THE COMPUTATIONAL TECHNOLOGY INTENDED FOR ANALYSIS AND RESTRUCTURING SOFTWARE TEXTS ON THE PRINCIPLES OF LOGIC-SEMANTIC ANALYSIS AND SYNTHESIS

The principles of the computational modeling technology, which is oriented to efficient analysis of unknown software and restructuring of the software text, are outlined. A software complex intended for semantic code analysis without execution of the program has been elaborated on the basis of this technology. The analytical technology presumes automatic constructing of graphs of relations between the functional-semantic clusters in the texts. The depth of analysis may vary from determination of separate operators' behavior (functions) to analysis of the total code, and its results may be used in time optimal finding errors in the program or verification of its correspondence to the specification. The technology of restructuring of the texts analyzed presumes reconstruction of the texts according to some algorithm and complementing the relation graphs with additional clusters.

**Keywords:** technologies of computational modeling; semantic analysis of software texts; restructuring software texts; static code analysis; relation graphs.

### Введение

В ходе технолого-коммерческой гонки второго десятилетия XXI века появляются новые высокопроизводительные вычислительные системы. Однако не секрет, что они строятся на старых принципах. Остаются нерешенными важнейшие проблемы, например, создание программных систем высокой надежности, что важнее, чем наличие систем высокой производительности, или вычислительных технологий, предполагающих быстрое создание и быстрое реструктурирование программных продуктов, что не менее важно. Их нерешенность постоянно сказывается и непременно станет существенным препятствием движению вперед в ближайшем будущем. В аспекте второй проблемы мы рассмотрим один из возможных подходов к решению задач оперативного смыслового анализа и реструктурирования готовых программ в целях быстрого создания новых программных продуктов на их основе. Мы будем говорить о разработанной нами технологии вычислительного моделирования, которая рассчитана на решение этих задач.

### 1. Постановка задачи.

Прежде всего, нужно сказать о причинах необходимости в указанной технологии. Создание программы – сложный и небыстрый процесс. Вот почему всегда существовал искуса прибегнуть к повторному использованию уже созданных программ или их блоков. На начало XXI века в национальном банке алгоритмов и программ (БАП) накоплено огромное количество талантливых разработок, которые легли “мертвым грузом”. Не меньшее их количество можно найти в WWW. Но вот – проблема: каждый раз, когда ставится задача создания новой программной системы, повторяется одна и та же история: программист надеется воспользоваться резервами БАП, он знает, что БАП содержит массу алгоритмов с открытым исходным кодом на C, C++, Java, Pascal и

других языках. Но, приступая к решению конкретной задачи, он вынужден вновь и вновь, как если бы все те базы программ были пусты, каждый раз программировать все с нуля. Оценки показывают, что так результат достигается быстрее.

Действительно, любой такой программный продукт “издали” выглядит как вроде бы вполне определенный и понятный, однако при ближайшем рассмотрении, “вблизи”, он оказывается совершенно иным. Несмотря на видимость открытости кода, анализ текста таких программ оказывается затруднительным до такой степени, как если бы их исходный код был закрыт. Это объясняется колоссальными трудностями (1) оптимального по времени содержательного анализа готовых программ, (2) анализа их интенционального (смыслового) предназначения согласно отсутствующей спецификации, а также (3) нерешенностью задач автоматического реструктурирования блоков готовых программ и встраивания их в новый программный продукт. Эти трудности и нужно было преодолеть.

## **2. О проблемах, препятствующих эффективному анализу готовых программ в целях их повторного использования**

Есть несколько объяснений невозможности быстрого и качественного анализа готовой программы: 1) реальная объектная программная система (ОПС) содержит тысячи файлов, сотни тысяч строк и десятки тысяч отношений между функциями; 2) неочевидны базовые принципы, на основе которых она построена; 3) последнее предполагает неочевидность базовых механизмов, поддерживающих технологию программирования, использованную при создании этой программы, и 4) невозможность быстрого построения модели этой ОПС.

К *принципиальным недостаткам* практически всех ОПС, которые доступны в БАП, следует отнести то, что: (1) они не построены (и сейчас все еще не строятся) на единых принципах модельной логико-смысловой организации. Говоря конкретно, программы из БАП (а) имеют различную, причем, часто неочевидную структуру, (б) содержат множество подпрограмм с неочевидной структурой и назначением, и, следовательно, каждая из этих подпрограмм не оказывается органичным и легко вычленимым элементом в модели ОПС. В итоге, анализ такой ОПС требует очень большого времени, т.к. будучи организованы лишь в структурно-логическом отношении, они, как правило, изначально не организованы в (а) логико-смысловом и (б) интенционально-смысловом (т.е. связанном с их назначением) отношениях, а потому очень трудно оценить пригодность ОПС и их подпрограмм для повторного использования во вновь разрабатываемой программе. Кроме того, (2) не ясно, обеспечит ли использование готовых модулей в создаваемой программе высокую эффективность и надежность программного продукта, если их погружение в новую программу будет осуществлено в соответствии с известными технологиями программирования.

Гарантии надежности создаваемых программных систем в перспективе требуют обеспечить их *прозрачность*, причем не только в структурном отношении, а, прежде всего, в логико-смысловом и интенционально-смысловом отношениях. Пока же представление любой программы в БАП является безмодельным, ограничено текстовой формой. Отражение даже общего информационно-логического содержания и назначения ОПС не предполагается. Между тем, (3) важно обеспечить не одну, а множество форм модельного представления алгоритма программы, которые были бы удобны для аналитика и пользователя, например, а) отображение функциональных логических, отношений и зависимостей в форме графов или сетей, б) отображение логико-смысловых отношений в терминах логических моделей или формул, в) модельное отображение интенционального смысла алгоритма и, как следствие, программы (последнее не оно и то же, что логические отношения внутри модуля или между модулями). Анализ множества таких отображений позволил бы упростить понимание логико-смысловой организации модулей в объектной программе и, следовательно, оценить их пригодность для целей повторного использования. Практика программирования свидетельствует о такой необходимости. В итоге, (4) пока логико-смысловую модель любой программной системы невозможно оптимально по времени распознать, следовательно, невозможно вычленивать из программы полезный элемент. Как следствие, (5) невозможно изменить или полностью перестроить объектную программную систему с учетом возникших потребностей пользователя, а проблемы с адаптацией программного продукта к потребностям пользователя возникают постоянно.

### 3. О проблемах, связанных с анализом интенционального смысла программы.

Проблемы, связанные с анализом интенционального смысла программы, заложенного в ее алгоритме, очень напоминают проблемы, связанные со сложностями смыслового анализа текстов на естественном языке (ЕЯ) [1]. Если внешняя форма ЕЯ-текста прозрачна, т.е. является экспликацией ясного, конкретного и логически выстроенного смыслового содержания, то такой текст легко поддается анализу, и его смысл может быть понят адресатом. Однако часто смысл ЕЯ-текста оказывается неочевидным, т.к. выражен: (а) невербально, т.е. не через логическую семантику конструкций текста, а средствами подтекста или экстралингвистики; (б) вербально, но неявно (напр. иносказательно). Тогда его очень трудно выявить путем анализа даже вроде бы очевидной, логически выстроенной вербальной последовательности текста. По существу, решение задач эффективного анализа и распознавания интенционального смысла частей текста программы и программы в целом (т.е. их назначения согласно спецификации) затрудняют аналогичные проблемы: 1) проблемы, связанные с неочевидностью логико-смысловых отношений в тексте; 2) проблема, связанная с тем, что старые программы писались на принципах старых технологий программирования (как если бы – старых форм ЕЯ). Вот почему необходимо добиваться того, чтобы все создаваемые программы строились на некоторых единых, универсальных принципах, сформулированных в [2].

### 4. Принципы анализа программ на основе технологии вычислительного моделирования

Основы технологии вычислительного моделирования (ТВМ) рассмотрены в [2]. Анализ текстов программ на основе ТВМ предполагает: (1) обязательное построение интегральной семантической модели (ИСМ) анализируемой программы и ИСМ вновь создаваемой программы; (2) разработку подхода к решению задачи по созданию таких моделей и выполнению моделирующего отображения; (3) создание средств эффективного представления функциональной семантики модели; (4) поиск способов упрощения анализа на моделях.

Один из наиболее эффективных подходов к задачам 1 и 2, решение которых предшествует решению задач анализа текстов, может быть моделирующее отображение программ. Но одно дело – моделирование текстов простых программ, и совсем другое дело – моделирование текстов больших программных систем, состоящих из сотен и тысяч файлов. Вот почему необходимо решение задачи 4, т.е. нужны способы (i) простого модельного представления текстовых объектов и (ii) упрощенного анализа текстов на моделях.

Что касается решения задачи 3, то моделирующее отображение текста программы должно (и может) передавать ее функционально-смысловую структуру, т.е. не отношений между блоками программы, а структуру отношений между функциями, выполняемыми этими блоками, учитывая, что эти функции отражают смысл, выражающийся в динамике отношений в исполняемой программе (Рис. 1). Очевидно, что данный подход применим лишь к текстам программ, написанных на языках, к которым применимо понятие *функция*.

Если не удастся выявить структуру отношений между всеми функциями текста ОПС, т.к. она велика, то необходимо как-то упростить анализ отношений. Этого можно добиться путем грамотной декомпозиции модели ОПС на функционально-смысловые (ФС-) блоки, по уровням: ФС-структура ОПС в целом – ФС-структура на уровне модулей ОПС – ФС-структура на уровне отдельных операторов ОПС.

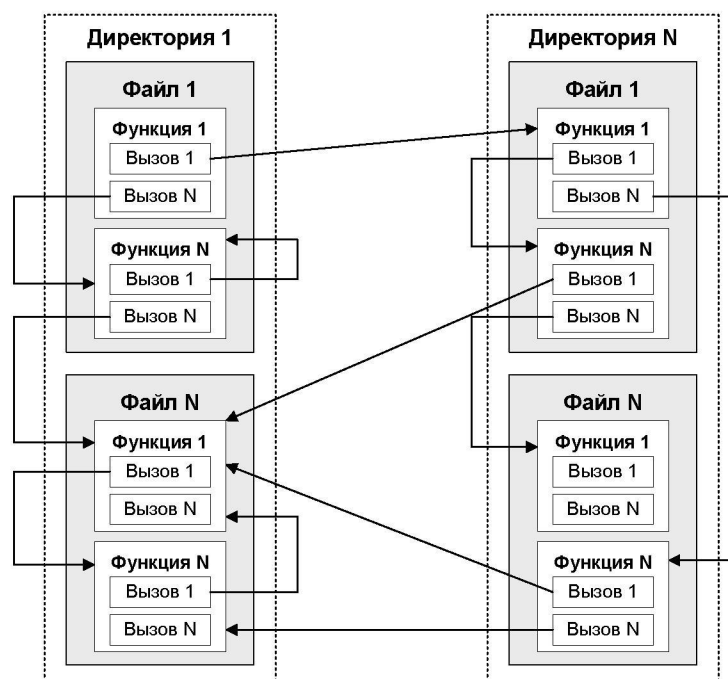


Рис. 1. Возможные отношения между функциями в модели анализируемой и моделируемой ОПС

Основу операций с множеством отношений в анализируемой ОПС составляют операции, предполагающие декомпозицию на модельные ФС-блоки.

Важно обеспечить многообразие способов моделирующего отображения модели (по сути – алгоритма) программы. Это могут быть, например, отображения в форме текстов, отображения в форме: (i) функциональных графов, (ii) сетей отношений, (iii) логических отношений в терминах логических формул, (iv) семантических отношений. Важно, чтобы аналитик имел возможность строить различные модельные отображения ОПС, их модулей и результатов вычислений, которые соответствовали бы растущим потребностям в многообразных формах представления программ и данных о них (Рис. 2).

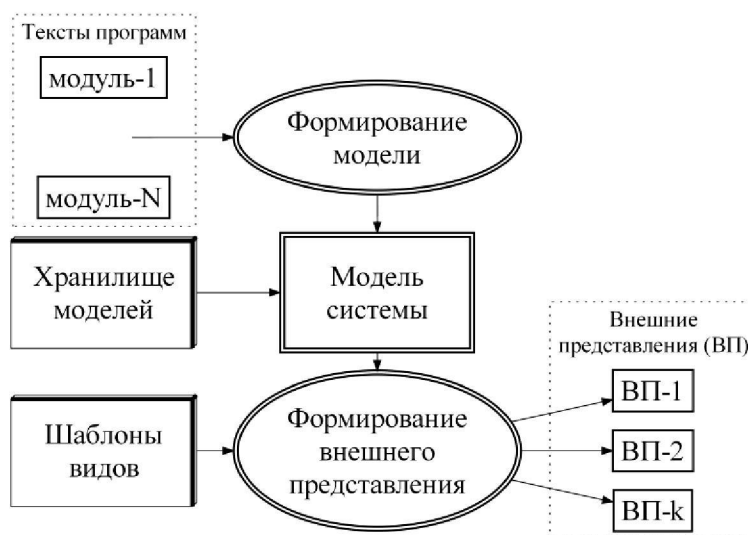


Рис. 2. Условное представление подхода к моделированию ОПС на основе технологии вычислительного моделирования

Одним из наиболее эффективных способов моделирующего отображения алгоритма программы может быть ее отображение, предполагающее формирование графа (или иерархии графов) функционально-смысловых отношений с различными уровнями представления внутренних



объектов и связей. Именно этот подход был использован нами при создании программного комплекса, предназначенного для содержательного анализа кода без выполнения программы.

### **5. Подход к реструктурированию программ**

Технология реструктурирования текстов программ допускает (i) перестройку текстов по заданному алгоритму и (ii) пополнение графов связей в тексте программы дополнительными кластерами. Принцип реструктурирования текстов программ на базе ТВМ основан на том, что модельная декомпозиция ОПС осуществляется на ФС-блоки. При необходимости ФС-блоки модели проще интегрируются обратно, в единое целое. Предполагается, что по готовым моделям программ (как прецедентов), накапливаемым в хранилище моделей программ, с учетом набора шаблонов, можно – как по задаваемым шаблонам – строить графы связей новых программ и, таким образом, формировать связи между различными формами внешнего представления программ и с различной их детализацией.

Важно то, что подход к анализу сложного открытого программного обеспечения (ПО) на основе ТВМ, предполагающей моделирование текстов программ на графах, многообразное отображение и использование вычислительных технологий анализа его логики, позволяет создавать программные системы, функциональное расширение которых (на базе открытого ПО и по разработанной технологии) может быть осуществлено за предельно короткие сроки.

### **6. О воплощении вычислительной технологии для исследования программных систем**

Существует множество подходов к решению поставленных нами задач. Мы выбрали тот, который предполагает создание практически работающей системы и практическое опробование на ней возможности эффективного исследования ОПС. Особенно сложно реализовать такой подход в случае, если разрабатываемая система должна выполнять функции логико-смыслового анализа. И все же авторы избрали именно этот путь.

На принципах, описанных выше, удалось построить эффективную технологию и воплотить ее в форме аналитического программного комплекса. Инструментальными основами технологии стали (1) универсальная инструментальная среда программирования ЗИРУС [3] и (2) подход, предполагавший представление программ в виде моделей данных и набора отношений. Он основан на общих принципах теории баз данных, применении теории графов в качестве аппарата модельного отображения [4] и рассчитан на формирование различных внешних текстовых и графических модельных отображений программ. Существенный сдвиг вперед был обеспечен за счет решения, связанного с упрощением представления сложных алгоритмов больших программ. На этой основе, в рамках проекта по формированию графов функциональных связей анализируемой программы по ее исходному коду, сначала был разработан аналитический программный комплекс АПК-1, предназначенный для автоматизированного анализа обширного класса модулей программ и программных систем в целях изучения возможности их повторного использования во вновь создаваемых программных продуктах. Логическая основа этой аналитической системы написана на языках Lua и Python. Система предполагает решение задачи статического анализа кода ОПС с помощью технологии автоматического построения графов связей в ее тексте, причем, анализ производится без выполнения программы, над версией исходного кода. Глубина анализа может варьировать от определения поведения отдельных операторов программы до анализа всего исходного кода.

Позднее, на основе этой аналитической системы, была разработана вычислительная система, предназначенная уже для исследования и модификации программных систем с четом итогов их логико-смыслового анализа. Эта версия предлагает новую возможность: модификация программы в соответствии с возникшими потребностями. Теперь в процессе анализа можно применять предустановленные шаблоны различных представлений графов, а также дополнять графы новыми кластерами, созданными аналитиком-пользователем или программистом-пользователем.

### **Заключение**

Итак, на основе технологии вычислительного моделирования (ТВМ) построен аналитический комплекс АПК-2. Он выполняет статический исходного кода ОПС и по шаблонам, задаваемым пользователям (а при обучении – по учебным шаблонам), строит графы связей различного внешнего представления и необходимой степени детализации (см. Рис. 3, где представлены итоги частичного анализа связей внутри исходного кода и связей с множеством функций типа sys).



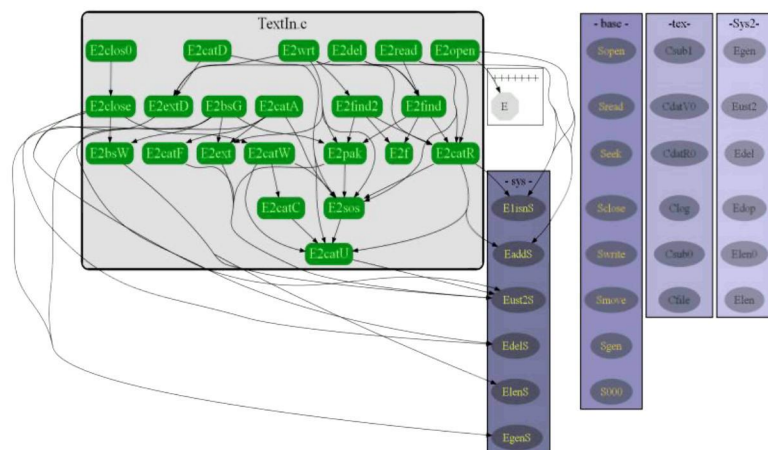


Рис. 3. Граф функциональных связей, построенный в итоге первого этапа анализа текста

Выполненный анализ позволяет аналитику, использующему АПК-2, оценить иерархию связей типа «Потомок → Родитель» в анализируемой ОПС, выявить неиспользуемые функциональные модули, рекурсии, возможные ошибки, идентифицировать свойства ОПС, проверить ее соответствие спецификации.

### Литература

1. Чернышов М.Ю. Проблема идентификации вербально выраженных мыслей-скрепов как медиаторов смысловой интегративности текстов. – М.: Наука и право, 2010. – 152 с.
2. Абасов Н.В., Чернышов М.Ю. На пути к вычислительной технологии для эффективного исследования программ и программных систем, основанной на принципах логико-смыслового анализа // Информационные технологии и высокопроизводительные вычисления: материалы междунар. конф. (Хабаровск, 4–6 октября 2011 г.). Хабаровск: Изд-во ТОГУ, 2011. С. 3–15.
3. Абасов Н.В. Основы универсальной среды программирования ЗИРУС // Вестник ИрГТУ. 2006. № 2(26). С. 62–68.
4. Чернышов М.Ю., Абасов Н.В. Основы вычислительной технологии, предназначенной для исследования программных систем и основанной на принципах логико-смыслового анализа // Вестник ТОГУ. 2012. № 2. С. 23–32.

*Чернышов Михаил Юрьевич* – канд. филол. наук, зав. научно-методической частью, Президиум Иркутского научного центра СО РАН, г. Иркутск, тел. моб. 8-914-932-74-90, e-mail: Michael\_Yu\_Chernyshov@mail.ru  
*Абасов Николай Викторович* – канд. техн. наук, ведущий научный сотрудник, Институт систем энергетики им. Л.А.Мелентьева СО РАН, г. Иркутск, e-mail: nva@isem.sei.irk.ru  
*Осипчук Евгений Николаевич* – аспирант, Институт систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск, e-mail: evgeny-osipchuk@yandex.ru

*Chernyshov Mikhail Yurievich* – candidate of philological sciences (mathematical linguistics, artificial intelligence), head of the research-methodological division, Presidium of Irkutsk Scientific Center, Siberian Branch of RAS, Irkutsk.

*Abasov Nikolay Viktorovich* – candidate of engineering sciences (computational technologies), leading researcher, Melentyev Institute of Energy Systems, Siberian Branch of RAS, Irkutsk.

*Osipchuk Evgeny Nikolayevich* – postgraduate, Melentyev Institute of Energy Systems, Siberian Branch of RAS, Irkutsk.

УДК 004.9:576.311.33:577.115.3:577.171.54

© *М.Ю. Чернышов, В.Н. Нурминский, Н.В. Абасов, Е.Н. Осипчук*

## МЕТОД КОЛИЧЕСТВЕННОГО АНАЛИЗА ОБЪЕКТОВ НА ВИДЕОИЗОБРАЖЕНИЯХ, РАССЧИТАННЫЙ НА ИСПОЛЬЗОВАНИЕ В ГИС-ТЕХНОЛОГИЯХ И АНАЛИЗЕ МИКРОСКОПИЧЕСКИХ ВИДЕОИЗОБРАЖЕНИЙ

Разработан метод количественного анализа состава видеоизображений удаленных макроскопических объектов в ГИС-технологиях и микроскопических видеоизображений клеточных препаратов в медико-биологических исследованиях. Его программная реализация предполагает применение аналитических процедур фильтрации, сегментации видеоизображения и вычитания фона. Метод воплощен в форме компьютерной программы.

**Ключевые слова:** метод количественного анализа, ГИС-технологии, микроскопические медико-биологические препараты, видеоизображение, сегментация видеоизображения, количественный анализ состава видеоизображения, программа количественного анализа.

*M.Yu. Chernyshov, V.N. Nurminsky, N.V. Abasov, E.N. Osipchuk*

## THE METHOD OF QUNTITATIVE ANALYSIS OF OBJECTS ON VISEOIMAGES, WHICH IS INTENDED FOR APPLICATIONS IN GIS-TECHNOLOGIES AND IN ANALYSIS OF MICROSCOPIC VISEOIMAGES

A method for quantitative analysis of video-images of distant macroscopic objects in GIS-technologies and microscopic video-images of cell preparations in medical-biological investigations has been elaborated. Its software implementation presumes application of analytical procedures of filtering, segmentation of the video-image and subtracting the background. The method has been implemented in the form of a software system.

**Keywords:** method of quantitative analysis, GIS-technologies, microscopic medical-biological preparations, video-image, video-image segmentation, quantitative analysis of video-image content, software for quantitative analysis.

### 1. Введение

Решение задач, стоящих сегодня перед наукой, требует привлечения инновационных методов и алгоритмов, позволяющих выполнять 1) качественный и количественный анализ объектов, наблюдаемых на Земле из космоса средствами ГИС-технологий (картография, анализ полезных ископаемых, мониторинг лесных пожаров и т.д.); 2) медико-биологическую морфометрию (структурный и количественный анализ) состава микрообъектов, развитие методов которой способствует получению средств качественной медицинской диагностики заболеваний; 3) исследование микроструктуры металлов и иных наноматериалов. Решение этих задач связано с качественным и количественным анализом растровых видеоизображений и требует разработки методов, алгоритмов и программ, обеспечивающих распознавание образов, программ компьютерной обработки, сегментации и последующего анализа растровых и цифровых изображений, алгоритмов обработки одномерных сигналов, статистической обработки данных. Известно множество подходов к решению этих задач, предложены алгоритмы обработки, сегментации и анализа изображений, но их нельзя признать простыми и надежными.

Например, согласно традиционной практике, прежде выполнения операций морфометрии, объекты на медико-биологических препаратах обычно выделяют, например, путем окрашивания с помощью специальных зондов-красителей. За счет окрашивания достигается цвето-яркостное различие объектов на окружающем фоне при облучении препарата лазером в полосе длин волн  $\lambda$  от  $10^{-5}$  до 1 нм. Но эта задача решается проще, если изображение содержит микрообъекты, которые близки по размеру и с точки зрения восприятия.

В работе описан метод морфометрии при минимальной априорной информации об объектах. Метод предполагает последовательное применение алгоритмов выделения поля объектов на растровом видеоизображении и оценки количества микроскопических объектов на нем.

## 2. Вычислительная технология количественного анализа растровых микроскопических видеоизображений

Для получения количественной информации об объектах на видеоизображении выполняется компьютерная обработка, предполагающая применение многопроцедурного алгоритма: 1) обработка видеоизображения; 2) фильтрация видеоизображения для устранения наводок и повышения контраста; 3) классификация объектов на изображении; 4) бинаризация; 5) количественный анализ числа объектов на видеоизображении.

**2.1. Многошаговый алгоритм количественного анализа растровых видеоизображений.** Изображения микрообъектов анализируемого препарата получали средствами цейтраферной компьютерной видеосъемки [2]. Установка ЦКВ, оборудованная микроскопом Биолам П-1 ( $\times 153$ ), позволяет получать изображения поверхности кюветы с пробой размером  $1820 \times 1365$  мкм, что соответствует размеру изображения в  $1600 \times 1200$  пикселей. При этом можно отслеживать динамику изменения количества клеток на последовательности изображений (рис. 1).

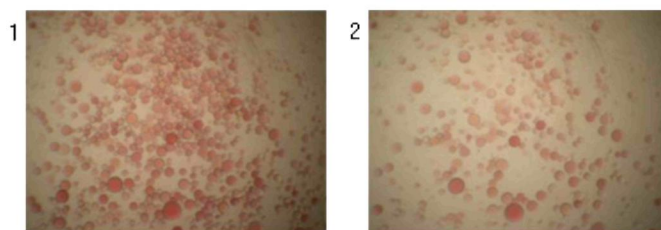


Рис. 1. Фрагмент последовательности изображений фракции изолированных вакуолей, полученной с помощью цейтраферной компьютерной видеосъемки

**Предварительная обработка видеоизображения.** Еще в исследовании В.О. Петрова, В.А. Камаева и С.В. Поройского [3] было выяснено, что в рассматриваемом случае традиционные подходы, предполагающие простой учет пространственных или статистических характеристик, не годятся. Этап предварительной обработки исходного видеоизображения предполагает улучшение яркости, контраста, т.е. различимости микрообъектов на нем. Пусть  $x(i, j) = x_{i,j}$  и  $y(i, j) = y_{i,j}$  – соответственно, значения яркости исходного изображения и изображения, получаемого на выходе после обработки, в пикселе кадра, имеющем декартовы координаты  $(i, j)$  (которые могут сохраняться в матрице как номера строк  $(i)$  и столбцов  $(j)$ ). Поэлементная обработка изображения подразумевает, что строится функциональная зависимость между значениями яркости такая, что  $y(i, j) = f_{ij}(x_{ij})$ , которая по значению исходного сигнала позволяет определить значение выходного сигнала. Очевидно, что значения функции  $f_{ij}(x_{i,j})$  зависят от координат, т.е., в общем случае, обработка кадра видеоизображения может быть неоднородной. Но для простоты в наших расчетах мы предполагали, что указанная зависимость между яркостями исходного и обработанного изображений является однородной, т.е. описывается функцией, значения которой одинаковы во всех точках кадра:  $y = kx(x_{\max} - x)$ , где  $x_{\max}$  – максимальное значение исходного сигнала;  $k$  – константа, позволяющая управлять динамическим диапазоном преобразования изображения. Функция  $y$ , описывающая данное преобразование, описывается параболой. Если  $y_{\max} = x_{\max}$ , то при  $k = 4/x_{\max}$  динамические диапазоны видеоизображений совпадают.

Задачу такой предварительной обработки видеоизображения можно решить, например, с помощью программы Adobe Photoshop. Но она не дает удовлетворительной декомпозиции изображения. Кроме того, микрообъекты имеют сфероидальную форму, и необходима фокусировка на “срезе микрообъектов посередине” (а не в их верхней или нижней части).

**Фильтрация видеоизображения** выполняется в целях устранения наводок, улучшения контраста, яркости и, следовательно, различимости микрообъектов. Был построен алгоритм выделения группы клеток для их подсчета (рис. 2).

Одна из проблем может быть связана с определением порога яркости  $m_{th}$ , с которым сравнивается яркость исходного видеоизображения и, таким образом, определяются значения яркости в каждой точке выходного изображения. Эффективным способом фильтрации изображения по яркости является Винеровская фильтрация. Подробное описание решения задачи фильтрации изображений с использованием оптимизационного подхода и решением уравнения Винера - Хопфа дано в статье М.С. Кириченко и С.В. Панина [1].

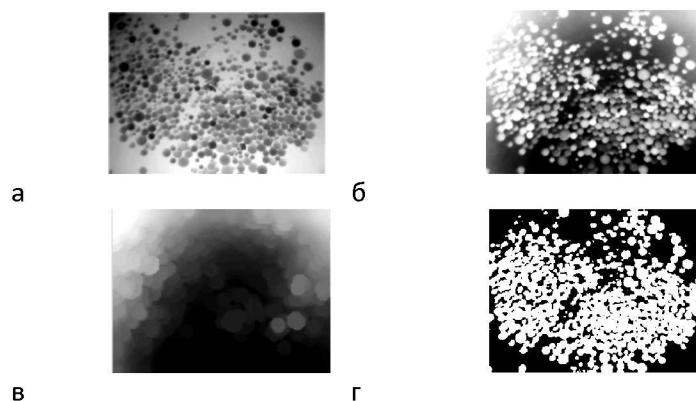


Рис. 2. Алгоритм выделения группы клеток для их подсчета (а – исходное изображение; б – итог первичной фильтрации; в – вычитаемый фон; г – результат бинаризации).

Решение этой задачи позволяет оценить результат линейной фильтрации полезного сигнала:

$$x^*(i, j) = \sum_{(i_1, j_1) \in S} \alpha(i_1, j_1) y(i - i_1, j - j_1),$$

где  $x^*(i, j) = x^*_{i,j}$  – результат фильтрации полезного сигнала  $x(i, j) = x_{i,j}$  в точке видеоизображения с координатами  $(i, j)$ ;  $S$  – множество координат точек, образующих окрестность точки изображения;  $\alpha(i_1, j_1)$  – весовые коэффициенты, совокупность которых дает двумерную импульсную характеристику фильтра. В свою очередь, это позволяет, например, оценить изменение средней яркости видеоизображения при его линейной фильтрации путем вычисления математического ожидания результата фильтрации полезного сигнала:

$$m_y^*(i, j) = m_x \sum_{(i_1, j_1) \in S} \alpha(i_1, j_1),$$

где предполагается, что средняя яркость  $m_x$  входного изображения  $x(\cdot)$  не зависит от координат, и, как следствие, средняя яркость  $m_x^*$  изображения  $y^*(\cdot)$  на выходе фильтра тоже постоянна во всех точках кадра. В ходе анализа важно поддерживать среднюю яркость кадра. Это достигается при выполнении равенства

$$\sum_{(i_1, j_1) \in S} \alpha(i_1, j_1) = 1.$$

**Задача бинаризации видеоизображения для выделения объектов по морфологическим (текстурным) особенностям** может быть решена путем его представления на языке теории множеств и соответствующих преобразований. Ее решение проще, если можно избежать сравнительного анализа цветов (объекты однородны по цвету). Тогда полутоновое изображение преобразуется в бинарное, и это все упрощает.

Поясним содержание процедуры бинаризации. Представим распределение яркости полутонового изображения на видеоизображении функцией  $f(x, y)$ , принимающей только неотрицательные значения, и предположим для простоты, что она имеет целые и равномерно разнесенные амплитуды яркости  $m = 0, 1, 2, \dots$ . Эта функция может быть реинтерпретирована как сумма бинарных функций  $f_b(i, j)$ , принимающих только пороговые (бинарные) значения, соответствующие исходному полутоновому видеоизображению, т.е.



$$f(x, y) = \sum_{(i, j) \in S} f_b(i, j),$$

где

$$f_b(i, j) = \begin{cases} 1 & \text{при } f(x, y) \geq m_{th}, \\ 0 & \text{при } f(x, y) < m_{th}, \end{cases}$$

Иными словами, каждая бинарная функция  $f_b(i, j)$  может быть представлена с помощью пороговых множеств  $X_b$  на плоскости, причем, каждая точка бинарного изображения:

$$\begin{aligned} (i, j) \in X_b, & \text{ если } f_b(i, j) = 1 \text{ и} \\ (i, j) \notin X_b, & \text{ если } f_b(i, j) = 0. \end{aligned}$$

В результате этого преобразования растровое полутоновое изображение преобразуется в бинарное (рис. 2г), которое соответствует матрице, содержащей только нули и единицы. Оно может быть сохранено в виде файла формата .txt, который затем используется на этапе количественного анализа, реализуемом с помощью специальной подпрограммы, вычисляющей относительную площадь изображения, занятую микрообъектами. Для решения задачи бинаризации может быть использована, например, система Mathcad 2000, преобразующая электронные изображения формата .bmp в матрицу, состоящую из чисел от 0 до 255, соответствующих пикселям монохромного видеоизображения. Пороговым ограничением амплитуды яркости  $m_{th}$  является некоторое число из указанного диапазона. В нашем случае значение этого порога было подобрано экспериментально и составило  $m_{th} = 24$ .

**Этап количественного анализа числа объектов на видеоизображении.** Решение этой задачи существенно упрощается, если микрообъекты имеют приблизительно равный диаметр, т.к. тогда достаточно вычислить площадь (в пикселях) выделенной области, занятой микрообъектами, и отнести ее к общей площади анализируемого видеоизображения.

Для определения количества  $M$  микрообъектов, идентифицируемых в полезном поле видеоизображения, осуществляется в два шага: (1) площадь полезного поля  $S_u$  определяется как количество пикселей  $n$  переднего плана (т.е. тех, которым соответствует значение  $f_b = 1$ ) на общем поле видеоизображения, содержащем  $N$  пикселей, т.е.  $S_u = n = \eta N$ ; (2)  $S_u$  делится на структурообразующий параметр  $\delta$ , соответствующий размеру объекта в пикселях:  $M = \eta N / \delta$ .

**2.2. Программная реализация метода.** Данный метод реализован как программный продукт в средах MATLAB (программа CellsPro) и Visual Studio C++ с использованием процедур обработки изображений библиотеки OpenCV (программа CellsProCV) [4].

### 3. Заключение

В целях экспериментально апробации алгоритма, на его основе были выполнены более 140 экспериментов с исследованием микроизображений клеточных вакуолей и эритроцитов крови в рамках программ исследований, проводимых в СИФИБР СО РАН. Результаты оказались весьма убедительными: (1) оценка вариабельности результатов, даваемых алгоритмом при повторных количественных анализах того же объекта составила  $\pm 5\%$ ; (2) изменения размеров видеоизображения в 2–4 раза практически не влияли на результаты количественного анализа; (3) анализ результатов работы алгоритма при изменении цвето-яркостных характеристик видеоизображения на 20–50% оказался удовлетворительным: вариабельность количественных результатов составила не более 5–15%.

Все сказанное позволяет утверждать, что программный комплекс, разработанный коллективом, руководимым М.С. Тарковым [4], предназначенный для количественного анализа микрообъектов на видеоизображениях, подходит для решения вышеуказанных задач.

### Литература

1. Кириченко М.С., Панин С.В. Разработка адаптивного алгоритма оценки информативности динамических признаков для обработки и анализа изображений // Вычислительные технологии. 2005. Т. 10. № 1. С. 58-70.
2. Стабильность вакуолярных мембран растений при осмотическом стрессе и воздействии редокс-агентов / В.Н. Нурминский и др. // Биологические мембраны. 2011. Т. 28, № 3. С. 224-229.
3. Петров В.О., Камаев В.А., Поройский С.В. Алгоритм текстурной сегментации растровых изображений при решении прикладных задач медико-биологического анализа // Современные проблемы науки и образования. 2009. № 6. С. 105-110.
4. Технология оценивания состава изображений клеточных препаратов для медико-биологических исследований / М.С. Тарков и др. // Математическое моделирование и вычислительно-информационные технологии в междисциплинарных научных исследованиях: тез. докл. всерос. конф. (Иркутск, 15–17 июня, 2011 г.). – Иркутск: Изд-во ИГЛУ, 2011. С. 109.

*Чернышов Михаил Юрьевич*, канд. филол. наук, почетный проф. Кембриджского ун-та (с 1991 г.), действительный иностранный член Британской королевской академии наук (Royal Society) (с 1998 г.), зав. научно-методической частью Президиума Иркутского научного центра СО РАН, г. Иркутск, тел. (3952) 425878. E-mail: Michael\_Yu\_Chernyshov@mail.ru.

*Нурминский Вадим Николаевич*, канд. биол. наук, старший научный сотрудник, Сибирский институт физиологии и биохимии растений СО РАН, г. Иркутск, сл.т. (3952) 425878. E-mail: cell@sifibr.irk.ru.

*Абасов Николай Викторович*, канд. техн. наук, ведущий научный сотрудник, Институт систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск. E-mail: nva@isem.sei.irk.ru.

*Осипчук Евгений Николаевич*, аспирант, Институт систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск. E-mail: evgeny-osipchuk@yandex.ru

*Chernyshov Mikhail Yurievich*, candidate of philological sciences (mathematical linguistics, artificial intelligence), head of the research-methodological division, Presidium of Irkutsk Scientific Center, Siberian Branch of RAS, Irkutsk.

*Nurminsky Vadim Nikolayevich*, candidate of biological sciences, senior researcher, Siberian Institute of Plant Physiology and Biochemistry Siberian Branch of RAS, Irkutsk, phone (3952) 425878. E-mail: cell@sifibr.irk.ru.

*Abasov Nikolay Viktorovich*, candidate of engineering sciences (computational technologies), leading researcher, Melentyev Institute of Energy Systems, Siberian Branch of RAS, Irkutsk.

*Osipchuk Evgeny Nickolayevich*, postgraduate, Melentyev Institute of Energy Systems, Siberian Branch of RAS, Irkutsk.

## ЛОГИКА НЕВЗАИМОДЕЙСТВУЮЩИХ ПРОГРАММ И РЕАКТИВНЫХ СИСТЕМ<sup>1</sup>

Логика программы – предикат (логическое утверждение), определяющий логику решения задачи и являющийся точным эквивалентом программы. Программа – это реализация логики решения в конструкциях языка программирования. Понятие логики программы обозначает также набор предикатов, истинных в разных точках программы. Эффективно и просто строится логика императивных и функциональных программ для различных выражений, операторов и типов, за исключением указателей. На базе логики программы определена формула тотальной корректности программы относительно спецификации в виде предусловия и постусловия. В целях дедуктивной верификации разработана система правил доказательства корректности программ для различных видов операторов.

Для реактивной системы логика программы – набор предикатов на переменных состояния программы. Предикат из этого набора истинен после исполнения некоторой очередной акции трассы, составленной перемешиванием акций параллельных взаимодействующих процессов. Акция – максимальный фрагмент кода программы процесса (без циклов внутри), для которого логика легко строится. Содержательное описание алгоритма функционирования реактивной системы формализуется в виде спецификации, представленной машиной конечных метасостояний как аппроксимации логики программы на множестве трасс. Разрабатываемый аппарат ориентирован на разработку, тестирование, моделирование и верификацию программной и аппаратной части встроенных систем аэрокосмической отрасли, энергетики, медицины и других приложений, где необходима предельная надежность систем.

**Ключевые слова:** логика программы, тотальная корректность программы, дедуктивная верификация, формальная семантика языка программирования, реактивная система.

V.I. Shelekhov, E.G. Tumurov

## LOGIC OF NON-INTERACTING PROGRAMS AND REACTIVE SYSTEMS

**Abstract.** The notion of a program logic is introduced to denote a set of predicates which are true in different program points. The program logic can be easily constructed for different kinds of statements of an imperative or a functional language with data types except pointers. For a non-interacting program, a total correctness formula based on the program logic is defined. The rules of program correctness proof have been developed for proving the statements of various kinds.

For a reactive system, the program logic is defined as a predicate on program state variables which is true after executing the current action of a trace constructed as a sequence of interleaved actions from different parallel interacting processes. An action is a maximal program fragment (without loops) of a process source code for which the program logic can be easily constructed. To analyze the behavior of a reactive system, the notion of a meta-state is introduced to denote an approximation of logic for a set of traces.

**Keywords:** program logic, total program correctness, deductive verification, formal semantics of a programming language, reactive system

### 1. Introduction

A *program logic* is a set of program properties expressed in the form of logical assertions (predicates) used by a programmer in the process of program development. These assertions originate from a mathematical problem solution and should be true in some points of the program execution; see [9] for details. More formally, the program logic at a program point is true for some values of the program variables if and only if there exists an execution reaching the point with those values.

The program logic differs from the specification of a program. It is the property of the program source code only, whereas specification is supplied by a programmer or customer. In deductive verification, the program logic can be extracted from a program with the help of formal semantics of a programming language and used in correctness condition formulas.

In Section 2, the notion of a program logic is defined for a subclass of programs not interacting with the external environment. This subclass includes programs for discrete and computational mathematical problems. The program logic can be easily constructed for different kinds of statements of an imperative or a functional language with data types except pointers. The logic of a loop statement can be defined as

<sup>1</sup> Работа выполнена при поддержке РФФИ, грант 12-01-00686.

the least fixed point of a recursive equation. The notion of (over)approximation of a program logic is defined. There are some interrelations between new notions and well-known ones. A loop invariant is a quite strong approximation of the program logic at the beginning of the loop. The program logic is exactly the strongest postcondition relative to the totally true precondition.

The method of deductive verification of non-interacting programs is described in Section 3. Formula (7) defines the total correctness of a program relative to its specification in the form of a precondition and a postcondition. To simplify generation of correctness formulas, specialization of formula (7) in the form of proving rules for different kinds of statements is introduced. The verification method is illustrated by an example program. An extended description of deductive verification with complete proofs of theorems, lemmas, and rules can be found in [9].

The condition of total correctness (7) is not valid for an interacting program, because the program may not finish or may have no resulting data and, as a consequence, will not have a postcondition. Nevertheless, the concept of a program logic may be extended over interacting programs. Section 4 introduces the first attempt to do it for the ordinary reactive systems which is only a subclass of various interacting programs.

The related work is presented in Section 5. An experience of deductive verification for about 30 small example programs is described in Conclusion.

## 2. Program logic for non-interacting programs

Let a program be the statement  $S(x, y)$  in an imperative or a functional language with the arguments  $x$  and results  $y$ , where  $x$  and  $y$  are variable sets. There are no interactions (inputs, outputs, etc.) between the program and its external environment. Thus  $S(x, y)$  implements a function from  $x$  to  $y$ .

The *program logic* is a predicate  $L(S(x, y))$  which is true for some values of the variables  $x$  and  $y$  if and only if there exists an execution reaching the end of the program  $S(x, y)$  with that values of variables  $x$  and  $y$ . The notion of program logic can be similarly defined for any construct (statement or expression) of a programming language. For example, the logic of the expression  $a/b$  is  $b \neq 0$  under the assumption that the language semantics does not restrict the representation of very large and small values.

Let us try to define the logic of different statements. Below, for modified variables of a statement, the primed variables correspond to the program state after executing the statement and unprimed ones represent the program state before executing the statement.

For the assignment statement  $a := E$ , where  $E$  is an expression, its logic  $L(a := E)$  is the predicate  $a' = E$ . Here,  $a$  and  $a'$  denote the initial and final values of the variable  $a$ , respectively.

Let us consider the sequential statement  $B; C$  defining the execution of the statement  $C$  after the execution of  $B$ . Let us define  $L(B; C)$  using the logics  $L(B)$  and  $L(C)$ . The case when the execution of  $C$  does not depend on the execution of  $B$  should be considered separately. We use the denotation  $A(x : y)$  for the statement  $A$  with the arguments  $x$  and results  $y$ .

Instead of the statement  $B; C$  we introduce *the superposition statement*  $B(x : z); C(z : y)$  and *the parallel statement*  $B(x : y) \parallel C(x : z)$ . The third statement in our consideration is the conditional statement **if** ( $E$ )  $B(x : y)$  **else**  $C(x : y)$ , where  $E$  is an expression that can depend on  $x$ . Under the assumptions that the variable sets  $x$ ,  $y$ , and  $z$  are disjoint and the set  $x$  can be empty, let us define the logics of the statements:

$$L(B(x : z); C(z : y)) \cong \exists z. L(B(x : z)) \& L(C(z : y)) \quad (1)$$

$$L(B(x : y) \parallel C(x : z)) \cong L(B(x : y)) \& L(C(x : z)) \quad (2)$$

$$L(\mathbf{if}(E)B(x : y)\mathbf{else}C(x : y)) \cong E \& L(B(x : y)) \vee \neg E \& L(C(x : y)) \quad (3)$$

The equivalent formula **if** ( $E$ )  $L(B(x : y))$  **else**  $L(C(x : y))$  can also be used for the logic of the conditional statement.

Below, for the statement  $B(u, x)$  with modified variables  $x$  and arguments  $u$ , we will also use the denotation of the form  $B(u, x : x')$ . Let the statement **while**  $E$  **do**  $B(u, x)$  **end** is denoted by  $W(u, x)$  with the results  $x$  and arguments  $u$  and  $x$ . It can be expressed in the form of the following recursive procedure:



**proc**  $W(u, x)$  { **if**  $(E)$  {  $B(u, x); W(u, x)$  } }

The logic of the **while** statement can be constructed by applying (3) and (1) to the procedure body and should be presented as the least fixed point of the following recursive equation:

$$L(W(u, x : x')) = \mathbf{if}(E) \exists z. L(B(u, x : z) \& L(W(u, z : x')) \mathbf{else} x' = x \quad (4)$$

Here,  $u$ ,  $x$ ,  $x'$ , and  $z$  are disjoint variable sets.

In a similar way, the program logic can be defined for each point of a program. The program logic  $L_p(S(z))$  of a program  $S$  at its point  $p$  is true for some values of the program variables  $z$  if and only if there exists an execution reaching the point  $p$  with the same values of  $z$ . At each point, the program logic can be constructed if there exists an algorithm constructing the logic of any statement.

$L_p(S)$  is the strongest formula which is true when the execution reaches the point  $p$ . So, in deductive verification, to establish the truth of a property  $X$  at the point  $p$  of the program  $S$ , it is sufficient to prove that the formula  $L_p(S) \Rightarrow X$  is true. To use a program logic  $L$  in deductive verification of programs, one should formally prove that the definition of the logic  $L$  is correct. First, the formal operational semantics should be developed for a programming language. Next, the definition of the program logic  $L$  should be formalized. For any statement in the programming language, one should prove that its program logic is consistent with the formal operational semantics.

For a point  $p$  in a program  $S$ , a predicate  $R$  is called an (over) *approximation of the logic*  $L_p(S)$  if  $L_p(S) \Rightarrow R$  holds. For a property  $X$  at the point  $p$ , if one can prove the formula  $R$  then obviously  $L_p(S) \Rightarrow X$ , and the property  $X$  is true. In deductive verification, one can use an approximation  $R$  if it is easier to construct than the logic  $L_p(S)$  or it has a simple structure allowing SMT solvers to be applied in proving. The approximation  $R$  should be strong enough in order for it to be possible to prove  $X$  from  $R$ . The predicate **true** is obviously an approximation of any program logic. But it is useless, because no property can be proved from it. However, the approximation **true** is safe and can be used as an initial state in constructing an appropriate approximation.

In Hoare deductive verification [2] of imperative programs, loop invariants are used in generating the verification condition to be proved. The automated construction of a loop invariant is a challenge. Note that a loop invariant is an approximation of the program logic at the beginning of the loop. This approximation should be strong enough in order that the conditions generated for loops be proved.

For a statement  $S$  and predicates  $P$  and  $Q$ , a *Hoare triple*  $\{P(x)\} S(x : y) \{Q(x, y)\}$  [2] is true if the following assertion holds: if  $P(x)$  is true before execution of  $S$  then  $Q(x, y)$  is true after it. The predicates  $P(x)$  and  $Q(x, y)$  are called a *precondition* and *postcondition*, respectively. The postcondition  $Q$  (denoted as  $sp(S, P)$ ) is called the *strongest postcondition* of  $S$  with respect to  $P$  if the triple  $\{P\} S \{Q\}$  is true and for all  $Q'$  such that  $\{P\} S \{Q'\}$  is true,  $Q \Rightarrow Q'$  holds. The program logic  $L(S(x : y))$  is exactly the predicate  $sp(S, \mathbf{true})$ .

### 3. Deductive verification of non-interacting programs

Let a program be a statement  $S(x : y)$ , where  $x$  and  $y$  are disjoint sets of variables. During the execution of the statement  $S(x : y)$ , no interactions with the external environment are possible. We suppose that all inputs take place before executing  $S(x : y)$  and all outputs are moved behind  $S(x : y)$ .

In deductive verification of a non-interacting program, it is sufficient to check that the specification of the program holds after program execution. The *specification* of the program  $S(x : y)$  can be represented in the form  $[P(x), Q(x, y)]$ , where the predicate  $P(x)$  is a precondition and  $Q(x, y)$  is a postcondition. *Correctness* of the program is defined by the following *correctness conditions*:

- the precondition  $P(x)$  should be true at the beginning of the program;
- the postcondition  $Q(x, y)$  should be true at the end of program;
- a program execution always terminates.

The third condition should hold because infinite execution of a non-interacting program is useless. It can be expressed by the predicate  $\exists y. L(S(x : y))$ , according to the definition of a program logic.

It is impossible to prove the truth of the precondition  $P(x)$  because  $P(x)$  does not depend on the program  $S(x : y)$ . Thus, it is sufficient to prove the second and third correctness conditions under the assumption that  $P(x)$  is true. The correctness conditions are expressed by the formulas:

$$P(x) \& L(S(x : y)) \Rightarrow Q(x, y) \quad (5)$$

$$P(x) \Rightarrow \exists y. L(S(x : y)) \quad (6)$$

Formula (5) is called a condition of *partial correctness*. Formula (6) is called a condition of *terminating a program*. Their conjunction defines the condition of *total correctness* of the program  $S(x : y)$  relative to the specification  $[P(x), Q(x, y)]$ :

$$\begin{aligned} \text{Corr}(S(x : y), P(x), Q(x, y)) &\equiv & (7) \\ P(x) \Rightarrow [L(S(x : y)) \Rightarrow Q(x, y)] \&\exists y. L(S(x : y)) \end{aligned}$$

Further the term “correctness” will be used in the sense of total correctness. Let us reformulate (7) as the following rule of program correctness:

$$\text{TC. } \frac{P(x) \& L(S(x : y)) \vdash Q(x, y); \quad P(x) \vdash \exists y. L(S(x : y))}{\text{Corr}(S(x : y), P(x), Q(x, y))}$$

The well-known notions of a total function and a single-valued function can be extended to predicates. The notions “the program  $S(x : y)$  is *single-valued*”, “the specification  $[P(x), Q(x, y)]$  is *total*”, and “the specification is *single-valued*” are defined by the following formulas:

$$P(x) \& L(S(x : y_1)) \& L(S(x : y_2)) \Rightarrow y_1 = y_2; \quad (8)$$

$$T(P(x), Q(x, y)) \equiv P(x) \Rightarrow \exists y. Q(x, y); \quad (9)$$

$$SV \ P \ x, Q \ x, y \equiv P \ x \ \& \ Q \ x, y_1 \ \& \ Q(x, y_2) \Rightarrow y_1 = y_2 \quad (10)$$

**Note.** A program totality property is precisely the correctness condition (6).

Non-single-valued program is nondeterministic. If the statement  $B(x : z)$  is not single-valued, it is possible for the superposition statement  $B(x : z); C(z : y)$  to terminate nondeterministically or not to terminate for some  $x$ . The notion of general correctness [8] has been developed for that case. The question arises if there exists a useful non-interacting program with this behavior in practice.

**Lemma 1.** If the program  $S(x : y)$  is correct relative to the specification  $[P(x), Q(x, y)]$ , then the specification is total.

The subformula  $L(S(x : y)) \Rightarrow Q(x, y)$  of the total correctness formula (7) defines the inference of the specification from the program logic. There are approaches (e.g. program synthesis) with a backward inference of a program from its specification<sup>2</sup>. The following theorem defines the conditions under which the program correctness can be proved by means of backward inference.

**Theorem 1 of identity between specification and program.**

Let the program  $S(x : y)$  be single-valued, the specification  $[P(x), Q(x, y)]$  be total, and the following formula be true:

$$P(x) \& Q(x, y) \Rightarrow L(S(x : y)) \quad (11)$$

Then the program  $S(x : y)$  is correct relative to the specification.

**Lemma 2.** Under the conditions of Theorem 1, the following formula holds:

$$P(x) \Rightarrow (L(S(x : y)) \equiv Q(x, y)) \quad .$$

**Corollary.** Under the conditions of Theorem 1, the specification  $[P(x), Q(x, y)]$  is single-valued.

If a specification is not single-valued, then formula (11) is false and therefore unprovable. Thus, formula (11) is applicable in verification for a single-valued specification only.

<sup>2</sup> In program synthesis, the program is extracted from a constructive proof of the theorem  $T(P(x), Q(x, y))$ .

The first condition of Theorem 1 is usually true. It is sufficient to prove that a program is single-valued if all basic operators used in the program are single-valued. This fact should be formally proved in the development of the formal operational semantics of a programming language.

Using the program logic for a programming language, we can construct automatically the correctness condition formula for the program  $S(x:y)$  on the base of formula (7) or (11). The resulting formula will be complicated and long even for short programs. Specialization of formulas (7) and (11) for different kinds of statements allows us to decompose a long correctness formula into a set of short and simple correctness formulas.

Below we present specialization of formula (7) in the form of the correctness proof rules for the superposition statement  $B(x:z);C(z:y)$ , parallel statement  $B(x:y) \parallel C(x:z)$ , and conditional statement  $\mathbf{if}(E)B(x:y) \mathbf{else} C(x:y)$  assuming the variable sets  $x$ ,  $y$ , and  $z$  are disjoint.

$$\begin{aligned}
 QP: & \frac{Corr(B(x:y), P(x), Q(x,y)); \quad Corr(C(x:z), P(x), R(x,z))}{Corr(B(x:y) \parallel C(x:z), P(x), Q(x,y) \ \& \ R(x,z))} \\
 QC: & \frac{Corr(B(x:y), P(x) \ \& \ E, Q(x,y)); Corr(C(x:y), P(x) \ \& \ \neg E, Q(x,y))}{Corr(\mathbf{if}(E)B(x:y) \mathbf{else} C(x:y), P(x), Q(x,y))} \\
 QS: & \frac{P(x) \vdash \exists z. L(B(x:z)); \quad Corr(C(z:y), P(x) \ \& \ L(B(x:z)), Q(x,y))}{Corr(B(x:z); C(z:y), P(x), Q(x,y))}
 \end{aligned}$$

Using these rules, one can reduce the correctness proof of superposition, parallel, and conditional statements to the correctness proofs of the substatements  $B$  and  $C$ . These rules can be repeatedly applied until the elementary statements are reached.

We also need to consider another case when there are specifications for the substatements  $B$  and  $C$ . The rules for that case are presented below.

$$\begin{aligned}
 RP: & \frac{Corr^*(B(x:y), P_B(x), Q_B(x,y)); \quad Corr^*(C(x:z), P_C(z), Q_C(x,z))}{P(x) \vdash P_B^*(x) \ \& \ P_C^*(x); \quad Q_B(x,y) \ \& \ Q_C(x,z) \vdash Q(x,y,z)} \\
 & \frac{Corr^*(B(x:z), P_B(x), Q_B(x,z)); \quad Corr^*(C(z:y), P_C(z), Q_C(z,y))}{RS: \quad \frac{P(x) \vdash P_B^*(x); \quad P(x) \ \& \ Q_B(x,z) \vdash P_C^*(z);}{P(x) \ \& \ Q_B(x,z) \ \& \ Q_C(z,y) \vdash Q(x,y)} \\
 & \frac{Corr^*(B(x:y), P_B(x), Q_B(x,y)); \quad Corr^*(C(x:y), P_C(x), Q_C(x,y));}{RC: \quad \frac{P(x) \ \& \ E \vdash P_B^*(x); \quad P(x) \ \& \ \neg E \vdash P_C^*(x);}{P(x) \ \& \ E \ \& \ Q_B(x,y) \vdash Q(x,y); \quad P(x) \ \& \ \neg E \ \& \ Q_C(x,y) \vdash Q(x,y)} \\
 & \frac{Corr^*(B(x:z), P_B(x), Q_B(x,z)); \quad Corr^*(C(z:y), P_C(z), Q_C(z,y));}{RB: \quad \frac{P(x) \vdash P_B^*(x) \ \& \ P_C^*(B(x)); \quad P(x) \ \& \ Q_C(B(x),y) \vdash Q(x,y)}{Corr(C(B(x):y), P(x), Q(x,y))}}
 \end{aligned}$$

In the above rules, if the substatement  $B$  is the recursive call of a procedure, the premise  $Corr^*(B, \dots)$  is omitted and  $P_B^*(x)$  is replaced by  $P_B(x) \ \& \ m(x) < m(z)$ , where  $z$  denotes the arguments of the recursive procedure and  $m$  is a *measure* function (of **nat** type) defined on arguments. If the substatement  $C$  is

a recursive call, the denotations  $Corr^*$  and  $P_C^*$  have the same meaning. When the substatement  $B$  (or  $C$ ) is not a recursive call,  $Corr^*$  and  $P^*$  simply denote  $Corr$  and  $P$ .

The last rule  $RB$  is a specialization of  $RS$ . Here,  $C(B(x):y)$  is equivalent to  $B(x:z); C(z:y)$ . Because  $z=B(x)$ ,  $z$  is replaced by  $B(x)$  in the rule  $RB$ .

The rules  $QS$  and  $T0$  include premises of two forms  $R(x,y) \Rightarrow \exists y.L(S(x:y))$  and  $R(x,y) \& L(S(x:y)) \Rightarrow H(x,y)$ , where  $R(x,y)$  and  $H(x,y)$  are arbitrary predicates. The next problem is to develop the rules for decomposing the occurrence  $L(S(x:y))$  in these forms. Let us present such rules for the superposition statement only.

$$FE: \frac{R(x) \vdash \exists z.L(B(x:z)); R(x) \& L(B(x:z)) \vdash \exists y.L(C(z:y))}{R(x) \vdash y.L(B(x:z); C(z:y))}$$

$$FL: \frac{L(B(x:z)) \& L(C(z:y)) \vdash H(x,y)}{L(B(x:z); C(z:y)) \vdash H(x,y)}$$

Let us illustrate application of the rules by an example program of multiplication of two natural numbers  $a$  and  $b$  with the result  $c$  using only the addition and subtraction operators. To obtain the algorithm in a tail-recursion form for transforming it automatically to an efficient **while** statement at the compilation stage, we need to consider a more general problem  $mult(a,b,d:c)$  with the specification  $[a \geq 0 \& b \geq 0 \& d \geq 0, c = a*b + d]$  and the following program:

$$\mathbf{proc} \mathbf{mult}(\mathbf{nat} a, b, d : \mathbf{nat} c) \tag{12}$$

$$\{ \mathbf{if} (a = 0) c = d \mathbf{else} \mathbf{mult}(a - 1, b, d + b : c) \} \mathbf{measure} a;$$

First, we introduce the definitions:

$$\mathbf{formula} P\mathbf{mult}(\mathbf{nat} a, b, d) = a \geq 0 \& b \geq 0 \& d \geq 0;$$

$$\mathbf{formula} Q\mathbf{mult}(\mathbf{nat} a, b, d, c) = c = a*b + d;$$

$$\mathbf{function} m(\mathbf{nat} a : \mathbf{nat}) = a;$$

Application of the rule  $QC$  to the program produces two formulas for proving:

$$Corr(c=d, P\mathbf{mult}(a,b,d) \& a=0, Q\mathbf{mult}(a,b,d,c))$$

$Corr(\mathbf{mult}(a-1, b, d+b:c), P\mathbf{mult}(a,b,d) \& a=0, Q\mathbf{mult}(a,b,d,c))$  Application of the rule  $T0$  to the first formula above produces the lemma:

$$\mathbf{lemma} P\mathbf{mult}(a,b,d) \& a=0 \& c=d \Rightarrow Q\mathbf{mult}(a,b,d,c);$$

To decompose the second formula above, let us replace the statement  $\mathbf{mult}(a-1, b, d+b:c)$  by an equivalent one

$|a', d'| = |a-1, d+b|; \mathbf{mult}(a', b, d' : c)$  and then apply the rule  $RB$ . Because the call of  $\mathbf{mult}$  is recursive, the rule  $RB$  is modified to

$$Corr(B(x:z), P_B(x), Q_B(x:z)); SV(P_B(x), Q_B(x,z));$$

$$P(x) \vdash P_B(x) \& P_C(B(x)) \& m(B(x)) < m(x);$$

$$RB^*: \frac{P(x) \& Q_C(B(x), y) \vdash Q(x, y)}{Corr(C(B(x):y), P(x), Q(x, y))}$$

Here  $P_B \equiv a-1 \geq 0$  and  $Q_B \equiv a'=a-1 \& d'=d+b$ . For the premises  $RB^*.3$  and  $RB^*.4$ , the following lemmas are generated:

$$\mathbf{lemma} P\mathbf{mult}(a,b,d) \& -a = 0 \Rightarrow a-1 \geq 0 \& m(a-1) < m(a)$$

$$\mathbf{lemma} P\mathbf{mult}(a,b,d) \& -a = 0 \& Q\mathbf{mult}(a-1, b, d+b, c)$$

$$\Rightarrow Q\mathbf{mult}(a, b, d, c)$$

The correctness of program (12) can be ensured by proving the above three lemmas.

One can compare the described verification of program (12) with the classic Hoare verification of the following imperative program obtained from program (12) by transforming the tail-recursion to the **while** statement:

$$c := d; \text{ while } a \neq 0 \text{ do } a := a-1; c := c+b \text{ end}$$

The superposition, parallel, and conditional statements can be used as the basis for construction of any programming language. The predicate programming language P [3], which is at the boundary of the functional and logical languages, has been developed on the base of these three kinds of statements. Its formal semantics, consisted of the operational and logical semantics, and the rules for proving the program correctness have been developed for the full language P [4, 6, 7]. An algorithm of generating the correctness condition formulas is implemented as a back-end in the compiler from P to C++.

An extended description of deductive verification with complete proofs of theorems, lemmas, and rules can be found in [9].

#### 4. Logic of reactive systems

The condition of total correctness (7) is not valid for an interacting program, because the program may not terminate or may have no resulting data and, as a consequence, will not have a postcondition. Nevertheless, the concept of a program logic can be extended over interacting programs. Our consideration is restricted within the set of ordinary reactive systems which is only a subset of various interacting programs. An illustration of notions defined below is given on an example of the Alternating Bit Protocol in Appendix A.

A reactive system consists of several processes executing in parallel and interacting among themselves and with the external environment by means of sending and receiving messages via unbounded FIFO channels. The following new kinds of statements are used in reactive systems:

**loop**{ $B$ } **send**  $m(e)$  **receive**  $m(x)$ { $B$ }**else**  $C$  **with** ( $y$ ){ $B$ }  $F \parallel G$  Here,  $B$  and  $C$  are statements,  $x$  and  $y$  are variable lists,  $e$  is an expression list,  $m$  is a message name,  $F \parallel G$  is a *parallel statement* for processes  $F$  and  $G$ , **with**( $y$ ){ $B$ } is a *protected statement*, blocking the access to variables  $y$  for other processes until terminating  $B$ .

When the statement **loop**{ $B$ } executes infinitely, the logic after the loop is **false**. For the case when the loop statement may terminate by executing the **break** statement inside  $B$ , it is problematically to define the logic of this loop adequately because the definition in the form the least fixed point like (4) would be wrong; an existence of other useful definitions is doubtful.

An *action* is a maximal program fragment of a source code not including loop statements. Logic of any action can be easily constructed using formulas (1)-(3). For a nested loop structure of the source code for a process, it is possible to construct an *action graph* where each edge defines *transition* from one action to another. A *trace* is an action sequence which can be produced by some path in the action graph. An execution of a process can be defined as the execution of the action sequence of some its trace.

For the parallel statement  $F \parallel G$ , a *state* is a set of variables with their values. The state includes global variables modified by either process. A channel between  $F$  and  $G$  belongs to the state for  $F \parallel G$ . We assume, the protected statements are inserted in a program so that the simultaneous access to each state variable is impossible for parallel processes. Note that a channel is implicitly protected while executing the **send** or **receive** statement for it. For actions  $a$  from  $F$  and  $b$  from  $G$ , the parallel composition  $a \mid b$  can be replaced by the sequential one either  $a; b$  or  $b; a$  producing the same modification of the state. This assertion is true under condition that for each action, none variable of the state can be modified more than once. In the case of multiple modifications within an action, the replacement of a parallel composition to the sequential one will be correct after applying the protected statements to some pieces of the action.

Let  $F_T$  and  $G_T$  be the sets of traces for the processes  $F$  and  $G$ , respectively. The set  $(F \parallel G)_T$  can be defined as the set of traces  $F_T * G_T$  constructed by all possible interleaving the traces from  $F_T$  and  $G_T$ . Now we can state the following assertion. A parallel execution of the statement  $F \parallel G$  is equivalent to the sequential execution of some its trace.

A *program logic of the parallel statement*  $F \parallel G$  for a trace from the set  $F_T * G_T$  is the program logic at the end of this trace. The trace is *reachable* if its logic is not **false**.

A specification of the parallel statement  $F \parallel G$  is a description of the behavior of its execution. Formal specification of  $F \parallel G$  can be given in the form of a *finite meta-state machine* (FMSM). A state (a node) of a FMSM is called a *meta-state*. It is attributed with the predicate on the state variables which

should be true when the execution of the FMSM reaches this meta-state. Below we just assume that the predicate is the value of this meta-state. A transition (an edge) in a FMSM is called a *meta-transition*. It consists of the sequence of the actions interleaved from  $F$  and  $G$ . A meta-transition may be supplied with a condition that should be true at the exit of the meta-transition.

The parallel statement  $F \parallel G$  should satisfy its FMSM specification. This correctness condition is expressed as the following assertion. For each reachable trace  $t$  for  $F \parallel G$ , there exists the path  $p_1, p_2, \dots, p_n$  in the FMSM, where  $p_j$  ( $j=1, \dots, n$ ) are meta-states, and there exist subtraces  $t_1, t_2, \dots, t_n$ , so that  $t = t_1 t_2 \dots t_n$  and for each  $k \leq n$ ,  $L(t_1 t_2 \dots t_k) P_k$ .

The FMSM specification can be used for detailed analysis of a reactive system. After proving that the specification correctly corresponds to the reactive system, it may be also used for proving the safety and liveness properties of the reactive system. The presented approach may be also used to develop a compiler back-end that automatically generates the correctness condition formulas for model checking and deductive verification of reactive systems.

### 5. Related work

The term “program logic” has been used as Hoare-style logic since 2003. Hoare logic defines the logic of a program implicitly by means of the rules for Hoare triples. The predicate transformers, like  $wp$ ,  $sp$ , etc, are other implicit definitions of the program logic. We use an explicit definition of the program logic as predicates that are true in some program points.

The notion of a program logic introduced in this paper coincides with Floyd’s interpretation of a program in his seminal paper [1]. An interpretation is a set of predicates at entries and exits of program constructs. In the flowchart language [1], each edge of a flowchart is supplied with a predicate. The language semantics is described by verification conditions for all kinds of constructs. Parameters of the verification condition for a construct are predicates at its entry and exit. According to the language semantics, the predicates of an interpretation should meet the verification conditions. For comparison, the program logic should be consistent with the formal operational semantics.

In verification based on the Hoare-style logic [2], only partial correctness conditions are verified whereas in our approach, total correctness of a program is defined by formula (7) based on the program logic. Another difference is that a measure function for a recursive procedure is easier to construct than an appropriate invariant for the loop obtained by replacing the tail-recursion.

The notion of a program logic is applicable in different fields of program verification. Some examples are presented below.

Logical formulas extracted from a program in the process of its reengineering are called a business logic. These formulas, of course, are an approximation of a program logic. The result of a static analysis of a program is some approximation of a program logic although usually it may be not represented in the form of logical formulas.

In the counterexample-guided abstraction refinement (CEGAR) method of software model checking [10, 11], abstract predicates, constructed for nodes of a program path to decide whether the path is reachable, are a lazy approximation of the program logic. An edge between two nodes corresponds to an elementary operator. In paper [11], an edge corresponds to a whole program fragment excluding loops called a large block. The large block encoding method [11] is based on constructing the program logic of large blocks used for calculation of abstract predicates.

Instead of extracting a logic from a program, logical formulas themselves can be considered as a program if they are executable. For the automatic proof system PVS [5], the PVS ground evaluator consists of a translator from an executable subset of the PVS specification language into Common Lisp [13]. The modeling, simulation and verification language MSVL for concurrent systems is an executable subset of the projection temporal logic [12].

The program logic differs from the specification of a program. It is the property of the program source code only, whereas specification is supplied by a programmer or customer. The program logic can be extracted from a program. But it is impossible to extract specification from the program source code as declared in [14, 15], especially for the case when specification is not single-valued. In fact, an extraction of the program logic is described in [14, 15]. In the best case, the program logic extracted can serve as some under-approximation of specification.



### Conclusion

The notion of a program logic is introduced for non-interacting programs and reactive systems to denote a set of predicates which are true in different program points. The program logic can be easily constructed for different kinds of statement of an imperative or a functional language with data types except pointers. For a non-interacting program, the total correctness formula (7), which uses the program logic, is defined. The rules of program correctness proof have been developed for proving the statements of various kinds. The notion of a program logic gives a proper understanding in various fields of program verification.

The superposition, parallel, and conditional statements can be used as the basis for constructing any programming language. The predicate programming language P [3], which is at the boundary of functional and logical languages, has been developed on the basis of these statements. Its formal semantics consisted of the operational and logical semantics has been developed for the full language P [4]. Based on formula (7), the proving rules have been developed for all constructs of the language P. On the basis of these rules, the detailed algorithm of generating the correctness condition formulas has been developed. It translates the correctness formulas to the language of the automatic proof system PVS [5]. It is implemented as a back-end in the compiler from P to C++.

The algorithm of generating the correctness formulas has been tested for about 30 small programs. The most valuable are Ling adder program [6] and the effective programs for the standard functions *floor*, *isqrt*, and *ilog2* [7]. The generated formulas for these programs have been proved in the PVS system. Proving the formulas on PVS was nontrivial and time-consuming. Twenty accidental errors were found in programs and specifications; only one error found in the *floor* program was dangerous.

The rules developed for proving the program correctness are also suitable for program synthesis [6, 7]. For the development of a program in the style of synthesis, the proof of the correctness formulas can be easier than in deductive verification.

Correctness formula (11) for the case of a single-valued specification looks simpler than general formula (7). Two years ago, the verification algorithm was developed on the basis of formula (11) only. Surprisingly, the proof of the correctness formulas appears to be more complicated and time-consuming than the proof of analogous formulas based on formula (7). The reason is that, along with the existence theorem  $T(P(x), Q(x, y))$ , we often need additionally to prove the unicity theorem. The important consequence of this fact is that the classic methods of program synthesis based on the existence theorem also suffer the same drawback.

### References

1. Floyd R.W. Assigning meaning to programs // Proc. Symposium in Applied Mathematics, 1967. – Vol. 19. P. 19-32.
2. Hoare C. A. R. An axiomatic basis for computer programming // Communications of the ACM. 1969. Vol. 12 (10). P. 576–585.
3. Karnaukhov N., Perchine D., and Shelekhov V. The predicate programming language P. – Preprint 153, Institute of Informatics Systems, Novosibirsk, 2009. – 44 p. (In Russian).
4. Shelekhov V. The language of calculus of computable predicates as a minimal kernel for functional languages // BULLETIN of the Novosibirsk Computing Center. Series: Computer Science. IIS Special Issue, 29, 2009. P. 107-117.
5. PVS Specification and Verification System. – SRI International. – URL: <http://pvs.csl.sri.com/>
6. Shelekhov V. Verification and synthesis of addition programs under the rules of statement correctness // Modeling and analysis of informations systems. – Yaroslavl, 2010. – Vol.17, No. 4. P. 101-110. (In Russian).
7. Shelekhov V. Verification and synthesis of effective programs for the standard functions *floor*, *isqrt*, and *ilog2* under predicate programming technology // Complex System: Control and Modelling Problems; XII International Conference. –Samara, 2010. P. 622-630. (In Russian).
8. Dawson J. E. Formalising General Correctnes // Electronic Notes in Theoretical Computer Science, Vol. 91, 2004. P. 21-42.
9. Shelekhov V. Rules of correctness proof for programs with simple logic // Modern Problems of Mathematics, Informatics and Bioinformatics, 2011. – 17 p. (in Russian). URL: [http://conf.nsc.ru/files/conferences/Lyap-100/fulltext/74974/75473/Shelekhov\\_prlogic.pdf](http://conf.nsc.ru/files/conferences/Lyap-100/fulltext/74974/75473/Shelekhov_prlogic.pdf)
10. Beyer D., Henzinger T., Jhala R., Majumdar R. The Software Model Checker Blast: Applications to Software Engineering // Int. Journal on Software Tools for Technology Transfer, 9, 2007. P. 505-525.
11. Beyer, D., Cimatti, A., Griggio, A., Keremoglu, M.E., Sebastiani, R. Software model checking via large-block encoding // Formal Methods in Computer-Aided Design, 2009. P. 25-32.
12. Mo D., Wang X., Duan Z. Asynchronous Communication in MSVL// ICFEM'2011, LNCS 6991, 2011. – P. 82-97.

13. Crow J., Owre S., Rushby J., Shankar N., and Stringer-Calvert D. Evaluating, Testing, and Animating PVS Specifications // CSL Tech. Report, 2001. – 55 p.
14. Dallmeier V., Knopp N., Mallon C., Hack S., Zeller A.: Generating test cases for specification mining // ISSTA '10, 2010. – P. 85-96.
15. Nguyen A. C., Khoo S.-C. Extracting Significant Specifications from Mining through Mutation Testing // ICFEM'2011, LNCS 6991, 2011. – P. 472-488.

*Шелехов Владимир Иванович*, канд. техн. наук, зав. лаб. системного программирования Института систем информатики, г. Новосибирск, тел. (383)330-27-21, e-mail: [vshel@iis.nsk.su](mailto:vshel@iis.nsk.su)

*Тумуров Эрдэм Гармаевич*, младший научный сотрудник Института систем информатики, г. Новосибирск, тел. (383)330-27-21, e-mail: [erdemus@gmail.com](mailto:erdemus@gmail.com)

*Vladimir Shelekhov*, candidate of technical sciences, laboratory head in A.P.Ershov Institute of Informatics Systems, Novosibirsk, Russia.

*Erdem Tumurov*, junior researcher in A.P.Ershov Institute of Informatics Systems, Novosibirsk, Russia.



### 3. Функциональный анализ и дифференциальные уравнения

УДК 519.676

©Т.А. Аверина

#### УСТОЙЧИВЫЕ ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ СТОХАСТИЧЕСКИХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В СМЫСЛЕ СТРАТОНОВИЧА<sup>1</sup>

В статье построены новые устойчивые численные методы решения стохастических дифференциальных уравнений в смысле Стратоновича.

**Ключевые слова:** стохастические дифференциальные уравнения, устойчивые численные методы.

*T.A. Averina*

#### NUMERICAL METHODS FOR SOLVING STOCHASTIC DIFFERENTIAL EQUATIONS IN THE STRAATONOVICH SENSE

In this article we construct new stable numerical methods for solving stochastic differential equations in the Straatonovich sense.

**Keywords:** stochastic differential equations, stable numerical methods.

#### Введение

Многие модели динамических систем в самых различных областях науки: радиотехнике, статистической механике, автоматическом управлении, химии, медицине, теории надежности и т.д., можно описать стохастическими дифференциальными уравнениями (СДУ). Сложность получаемых моделей затрудняет аналитическое исследование решений таких систем. Кроме того, для получения некоторых вероятностных характеристик решения, необходимых на практике, слабо развиты аналитические методы. В этих условиях на первый план стали выходить численные методы.

Как отмечено в работе [1], многие физические задачи, связанные с анализом быстропротекающих процессов в сильнонеравновесных средах, таких как термоядерная, лазерная, газоразрядная и космическая плазма, также можно описать с помощью СДУ. Причем, предельный переход к модели корректен только для СДУ в смысле Стратоновича. Актуальность построения устойчивых методов решения СДУ в смысле Стратоновича обсуждается в работе [2]. В работе [3] было предложено семейство численных методов для решения СДУ в смысле Стратоновича. В данной статье построен устойчивый численный метод из этого семейства. Построенный метод имеет 2-й порядок среднеквадратической сходимости для систем СДУ с одним шумом и 1-й порядок – в общем случае. Построенный метод рекомендуется для решения задач физики плазмы.

#### 1. Семейство численных методов решения СДУ в смысле Стратоновича

Пусть на вероятностном пространстве  $(\Omega, F, P)$  заданы: поток  $\sigma$ -алгебр  $\{F_t\}_{t \in [0, T]}$ ;  $n_w$ -мерный стандартный винеровский процесс  $w(t)$ ,  $t \in [0, T]$ , согласованный с  $\{F_t\}_{t \in [0, T]}$ , приращения которого  $w(t+s) - w(t)$  при  $s > 0$  не зависят от  $\sigma$ -алгебры  $F_t$ ;  $F_0$  - измеримый  $n_y$ -мерный случайный вектор  $y_0$ , независимый с  $w(t)$  при  $t \geq 0$ , причем  $E(|y_0|^2) < \infty$ .

Задача Коши для системы СДУ в смысле Стратоновича ставится следующим образом: найти  $n_y$ -мерный случайный процесс  $y(t)$ , для которого

$$y(t) = y_0 + \int_0^t f(y(\tau)) d\tau + \int_0^t \sigma(y(\tau)) \circ dw(\tau), y(0) = y_0, t \in [0, T], \quad (1)$$

где  $f(y(t))$  – измеримая по совокупности переменных  $n_y$ -мерная вектор-функция,  $\sigma(y(t))$  – измеримая по совокупности переменных матричная функция размера  $n_y \times n_w$ .

Для статистического моделирования траекторий решения систем СДУ в смысле Стратоновича (1) будем использовать семейство численных методов [3] вида

<sup>1</sup> Работа выполнена при финансовой поддержке грантов РФФИ (проекты № 11-01-00282 и № 12-01-00490)

$$\begin{aligned}
 y_{n+1} &= y_n + p_1 k_1 + p_2 k_2 + \sqrt{h}(q_{11}G_0 + q_{12}G_1 + q_{13}G_2)\zeta_n, \\
 k_1 &= \left( I - ha \frac{\partial f}{\partial y} \right)^{-1} \left( hf + q_1 \sqrt{h}G_0 \zeta_n + \frac{q_2 h}{2} \frac{\partial \sigma}{\partial y} \sigma \zeta_n^2 \right), \\
 G_0 &= \sigma(y_n), \quad G_1 = \sigma(y_n + \alpha_1 k_1 + q_3 \sqrt{h}G_0 \zeta_n + \frac{q_4 h}{2} \frac{\partial \sigma}{\partial y} \sigma \zeta_n^2), \\
 k_2 &= \left( I - ha \frac{\partial f}{\partial y} \right)^{-1} \left[ hf(y_n + \alpha_2 k_1 + q_5 \sqrt{h}G_0 \zeta_n + \frac{q_6 h}{2} \frac{\partial \sigma}{\partial y} \sigma \zeta_n^2) + q_7 \sqrt{h}G_1 \zeta_n + \frac{q_8 h}{2} \frac{\partial \sigma}{\partial y} \sigma \zeta_n^2 \right], \\
 G_2 &= \sigma(y_n + \alpha_3 k_1 + \alpha_4 k_2 + q_9 \sqrt{h}G_1 \zeta_n + \frac{q_{10} h}{2} \frac{\partial \sigma}{\partial y} \sigma \zeta_n^2).
 \end{aligned}
 \tag{2}$$

Здесь все функции, у которых не указан аргумент, вычисляются в точке  $y_n$ , где  $y_n$  – значения приближенного решения системы СДУ (1) в узлах сетки по времени  $t_n$ ;  $h$  – шаг интегрирования;  $a, p_i, q_i, \alpha_i$  – вещественные параметры метода;  $\zeta_n$  –  $n_w$ -мерный вектор независимых гауссовских случайных величин с нулевым математическим ожиданием и единичной дисперсией;  $\zeta_n$  – независимы с  $y_n$ . Данное семейство является обобщением методов Розенброка на случай стохастических дифференциальных уравнений. Методы Розенброка являются А-устойчивыми. Наличие параметра  $a$  в обобщенных методах Розенброка также позволяет улучшать их свойства устойчивости. Приведем необходимые определения. Для простоты полагаем, что  $h = t_{n+1} - t_n = T/K$ ;  $n = 0, \dots, K-1$ ;  $t_0 = 0$ .

**Определение 1** [4]. Численный метод *аппроксимирует* точное решение задачи Коши для СДУ (1) с порядком  $p$  в среднеквадратическом смысле, если для условного математического ожидания выполняется равенство

$$\max_{0 \leq n \leq K-1} E(|y(t_{n+1}) - y_{n+1}|^2 / y_n = y(t_n)) = O(h^{p+1}), \quad h \rightarrow 0.$$

**Определение 2** [4]. Численный метод *сходится с порядком  $p$  в среднеквадратическом смысле*, если

$$\max_{1 \leq n \leq K} E(|y(t_n) - y_n|^2 / y_0 = y(t_0)) = O(h^p), \quad h \rightarrow 0.$$

Как показано в [4], численный метод решения СДУ имеет  $p$ -й порядок сходимости в среднеквадратическом смысле, если он аппроксимирует точное решение с порядком  $p$  в среднеквадратическом смысле и также аппроксимирует математическое ожидание с порядком  $q \geq p/2$ .

**Определение 3** [4, 5]. Численный метод называется *асимптотически несмещенным (или устойчивым)* с шагом  $h > 0$ , если при его применении с этим шагом к скалярному линейному СДУ

$$y(t) = y_0 - \alpha \int_0^t y(\tau) d\tau + \int_0^t \sigma \circ dw(\tau), y(0) = y_0, t \in [0, T],
 \tag{3}$$

где  $\alpha, \sigma$  – вещественные коэффициенты,  $\alpha > 0$ , распределение численного решения  $y_n$  при  $n \rightarrow \infty$  сходится к нормальному распределению с нулевым математическим ожиданием и дисперсией  $\sigma^2 / (2\alpha)$ .

**Определение 4** [4, 5]. Интервал  $(x_0, 0)$  называется *интервалом асимптотической несмещенности (или устойчивости) метода*, если метод является асимптотически несмещенным с любым шагом  $h > 0$ , для которого  $-ch \in (x_0, 0)$ .

**Определение 5** [4, 5]. Численный метод называется  *$\gamma$ -асимптотически смещенным (или  $\gamma$ -устойчивым)* при  $0 < h \leq h_0$ , если при его применении с фиксированным шагом  $h$  к скалярному СДУ (3) с  $\alpha > 0$  распределение численного решения  $y_n$  при  $n \rightarrow \infty$  сходится к нормальному распределению с нулевым математическим ожиданием и дисперсией  $\bar{d}$ , причем

$$\left| \frac{\bar{d}}{\sigma^2 / 2\alpha} - 1 \right| \leq \gamma.$$

Понятие асимптотической несмещенности (или устойчивости) метода можно определить и для систем СДУ [6].

## 2. Асимптотическая несмещенность, аппроксимация и сходимость методов

**Теорема 1.** Семейство численных методов (2) содержит асимптотически несмещенные и  $\gamma$ -асимптотически смещенные численные методы. Асимптотически несмещенные численные методы из семейства (2) могут иметь только бесконечный интервал асимптотической несмещенности, и их параметры должны удовлетворять уравнениям

$$\alpha\beta = p_2q_5, \quad p_1 + p_2 = 1, \quad \alpha_2p_2 = 0, \quad a = 1/2. \quad (4)$$

При

$$\alpha\beta = p_2q_5, \quad p_1 + p_2 = 1, \quad a(p_1 + p_2) + \alpha_2p_2 = 1/2, \quad 1/4 \leq a \leq 1/2 \quad (5)$$

для любого сколь угодно малого  $\gamma$  возможно построение  $\gamma$ -асимптотически смещенного метода. Интервал  $\gamma$ -асимптотической смещенности будет определяться из неравенства:

$$\left| \frac{[z(a - \alpha_2p_2q_1) + 1]^2}{[z^2(a - 1/2)^2 + 2z(a - 1/4) + 1][2z(a - 1/4) + 1]} - 1 \right| \leq \gamma.$$

**Доказательство** теоремы основано на применении метода (2) к уравнению (3) и сравнении вероятностных характеристик полученного решения и точного решения (3).

Теперь найдем условия, при которых численный метод из семейства численных методов (2) сходится в среднеквадратическом смысле.

**Теорема 2.** Семейство численных методов (2) содержит подмножество численных методов, имеющих первый порядок среднеквадратической сходимости для произвольных систем СДУ и второй порядок для систем СДУ с одним шумом, а также в случае систем СДУ с постоянной матрицей  $\sigma$ . Параметры этих методов удовлетворяют уравнениям

$$p_1 + p_2 = 1, \quad p_1q_1 + p_2q_7 + q_{11} + q_{12} + q_{13} = 1, \quad (6)$$

$$\frac{p_1q_2 + p_2q_8}{2} + (\alpha_1q_1 + q_3)(p_2q_7 + q_{12}) + q_{13}(\alpha_3q_1 + \alpha_4q_7 + q_9) = \frac{1}{2}. \quad (7)$$

**Доказательство** теоремы основано на сравнении разложений точного и численного решений в ряд Тейлора и применении теоремы сходимости [4].

### Построенные численные методы решения СДУ

В работе [3] приведены некоторые численные методы решения СДУ в смысле Стратоновича из семейства (2).

Выделим из семейства методов (2) подсемейство

$$y_{n+1} = y_n + p_1k_1 + p_2k_2 + \sqrt{h}q(\sigma(y_{n+1}^p) + \sigma(y_n^p))\zeta_n, \quad (8)$$

$$k_1 = \left( I - ha \frac{\partial f}{\partial y} \right)^{-1} (hf(y_n) + \sqrt{h}\sigma(y_n)\zeta_n),$$

$$k_2 = \left( I - ha \frac{\partial f}{\partial y} \right)^{-1} [hf(y_{n+1}^p) + \sqrt{h}\sigma(y_{n+1}^p)\zeta_n], \quad y_{n+1}^p = y_n + sk_1$$

с пятью параметрами. В записи методов (8) учтена возможность построения  $\gamma$ -асимптотически смещенного метода. Параметры методов выберем так, чтобы метод имел 1-й порядок аппроксимации первых двух моментов в общем случае и 2-й – для систем с одним шумом. Приведем три таких метода с параметрами:

$$a = 3/8, \quad p_2 = 1/8, \quad p_1 = 7/8, \quad q = 3/8, \quad s = 1;$$

$$a = 1/3, \quad p_2 = 1/6, \quad p_1 = 5/6, \quad q = 1/3, \quad s = 1;$$

$$a = 1/4, \quad p_2 = 1/4, \quad p_1 = 3/4, \quad q = 1/4, \quad s = 1.$$

Если параметры численного метода удовлетворяют (4), (6) и (7), то это асимптотически несме-

шенный метод с 1-м порядком среднеквадратической сходимости для произвольных систем СДУ и 2-м - для систем СДУ с одним шумом, а также в случае систем СДУ с постоянной матрицей  $\sigma$ . Приведем пример такого метода:

$$y_{n+1} = y_n + \left( I - \frac{h}{2} \frac{\partial f}{\partial y} \right)^{-1} \frac{h}{2} (f(y_n) + f(y_{n+1}^p)) + \frac{\sqrt{h}}{2} (\sigma(y_{n+1}^p) + \sigma(y_n^p)) \zeta_n, \quad y_{n+1}^p = y_n + \sqrt{h} \sigma(y_n) \zeta_n.$$

Этот асимптотически несмещенный метод не требует вычисления производной матрицы  $\sigma$  и рекомендуется, в частности, для решения задач, рассматриваемых в [1, 2].

#### Литература

1. Змиевская Г.И. Стохастические аналоги неравновесных столкновительных процессов // Физика плазмы. 1997. Т. 23. № 4. С. 368-382.
2. Змиевская Г.И., Бондарева А.Л. Островки тонкой пленки полупроводника и численный эксперимент // Поверхность. Рентгеновские, синхротронные и нейтронные исследования. 2010. № 10. С. 50-58.
3. Аверина Т.А., Артемьев С.С. Новое семейство численных методов решения стохастических дифференциальных уравнений // Докл. АН СССР. 1986. Т. 288. № 4. С. 777-780.
4. Артемьев С.С. Численные методы решения задачи Коши для систем обыкновенных и стохастических дифференциальных уравнений. – Новосибирск: Изд. ВЦ СО РАН, 1993 – 156 с.
5. Артемьев С.С. Численное решение обыкновенных и стохастических дифференциальных уравнений. - Новосибирск: НГУ, 1995 – 103 с.
6. Аверина Т.А., Артемьев С.С., Некоторые вопросы построения и использования численных методов для решения систем стохастических дифференциальных уравнений. Новосибирск, 1987. (Препринт / АН СССР. Сиб. отд-ние. ВЦ; 728).

*Аверина Татьяна Александровна*, канд. физ.-мат. наук, старший научный сотрудник Института вычислительной математики и математической геофизики СО РАН, пр. акад. Лаврентьева, 6, г. Новосибирск, 630090, Россия; тел. (+7 383)330-77-21, e-mail: [ata@osmf.sccc.ru](mailto:ata@osmf.sccc.ru).

*Averina Tatyana Aleksandrovna*, candidate of physical and mathematical sciences, Senior Scientist of Institute of Computational Mathematics and Mathematical Geophysics SB RAS, pr. Akademika Lavrentjeva, 6, Novosibirsk, 630090, Russia; ph. (+7 383)330-77-21, E-mail: [ata@osmf.sccc.ru](mailto:ata@osmf.sccc.ru).

## ОБ ОДНОМ ОБОБЩЕНИИ ДЛЯ РЕШЕНИЯ НАЧАЛЬНО-КРАЕВОЙ ЗАДАЧИ О КОЛЕБАНИЯХ ПРОИЗВОЛЬНОГО ЧИСЛА ОСЦИЛЛЯТОРОВ НА СТЕРЖНЕ<sup>1</sup>

В статье обобщается решение начально-краевой задачи на случай с произвольным числом осцилляторов на стержне. Концы стержня имеют жесткое закрепление. Выводится условие ортогональности собственных форм колебаний. Решение ищется в виде разложения в ряд Фурье по собственным формам колебаний системы.

**Ключевые слова:** собственные частоты, собственные формы, условие ортогональности, начально-краевая задача, ряд Фурье.

*S.G. Barguev, T.V. Ausheev, A.D. Mizhidon*

## A GENERALIZATION FOR SOLVING INITIAL-BOUNDARY PROBLEM ABOUT FLUCTUATIONS OF ARBITRARY NUMBER OSCILLATORS ON THE ELASTIC CORE

In the article is the generalization of the solution of the initial-boundary value problem to the case of arbitrary number oscillators on the elastic core with fastened ends. Are obtained the orthogonally condition of the own forms fluctuation. The decision are fined in view resolving at Fourier series on the own fluctuation forms of the mechanical system. Are make the convergent ion analysis of obtained resolving.

Key words: generalization ,own frequency, own forms, orthogonally condition, initial-boundary problem, Fourier series .

### Введение

Отметим важную особенность при рассмотрении начально-краевой задачи, если на стержне имеются два и более осцилляторов: колебательный процесс в этом случае меняется качественно, так как осцилляторы начинают взаимодействовать не только со стержнем, но и между собой.

Тогда можно ставить задачу о гашении колебаний части осцилляторов, оказывая влияние на перераспределение энергии между ними путем подбора начальных условий.

### 1. Получение условия ортогональности

Рассмотрим механическую систему, представляющую собой однородный упругий стержень длины  $\ell$ , плотности  $\rho$ , модулем упругости  $E$ , моментом инерции  $J$  поперечного сечения стержня относительно нейтральной оси сечения, перпендикулярной плоскости колебаний, с закрепленными на нем в точках с абсциссами  $a_k$  телами массами  $m_k$ ,  $k = 1, 2, \dots, n$ , посредством пружин с жесткостями  $c_k$ ,  $k = 1, 2, \dots, n$ . Концы стержня жестко закреплены. Массы  $m_i$  могут перемещаться вертикально в направлении осей  $z_k$ . Колебания масс характеризуются функциями  $z_i(t)$ , перемещения точек стержня описываются функцией  $u(x, t)$ .

Гибридная система дифференциальных уравнений, описывающая движение механической системы имеет вид [1]

$$\begin{aligned} \frac{d^2 z_k}{dt^2} + p_k^2 (z_k - u(a_k, t)) &= 0 \\ \frac{\partial^2 u}{\partial t^2} + b \frac{\partial^4 u}{\partial x^4} &= \sum_{k=1}^n e_k (z_k - u(x, t)) \delta(x - a_k), \end{aligned} \quad (1)$$

где  $p_k^2 = \frac{c_k}{m_k}$ ,  $b = \frac{EJ}{\rho F}$ ,  $e_k = \frac{c_k}{\rho F}$ .

Краевые условия на концах стержня:

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проект 12-08-00309-А)

$$\begin{aligned} u(0, t) = u(l, t) = 0, \\ \frac{\partial u}{\partial x}(0, t) = \frac{\partial u}{\partial x}(l, t) = 0 \end{aligned} \tag{2}$$

Представив  $z_k(t), u(x, t)$  в виде  $z_k(t) = A_k \sin(\omega t + \alpha_k), u(x, t) = V(x) \sin(\omega t + \beta)$ , подставив в (1), после преобразований получаем

$$\begin{aligned} -\omega^2 A_k + p_k^2 (A_k - V(a_k)) = 0 \\ -\omega^2 V(x) + b \frac{d^4 V(x)}{dx^4} = \sum_{k=1}^n e_k (A_k - V(x)) \delta(x - a_k) \end{aligned} \tag{3}$$

с краевыми условиями на концах стержня

$$\begin{aligned} V(0) = V(l) = 0 \\ \frac{dV}{dx}(0) = \frac{dV}{dx}(l) = 0 \end{aligned} \tag{4}$$

Запишем (3) для частот  $\omega_i$  и  $\omega_j$ :

$$\begin{aligned} -\omega_i^2 A_{ki} + p_k^2 (A_{ki} - V_i(a_k)) = 0 \\ -\omega_i^2 V_i(x) + b \frac{d^4 V_i(x)}{dx^4} = \sum_{k=1}^n e_k (A_{ki} - V_i(x)) \delta(x - a_k) \end{aligned} \tag{5}$$

$$\begin{aligned} -\omega_j^2 A_{kj} + p_k^2 (A_{kj} - V_j(a_k)) = 0 \\ -\omega_j^2 V_j(x) + b \frac{d^4 V_j(x)}{dx^4} = \sum_{k=1}^n e_k (A_{kj} - V_j(x)) \delta(x - a_k) \end{aligned} \tag{6}$$

Из (5) и (6) получаем выражения  $A_{ki}, A_{kj}$  вида

$$A_{ki} = \frac{p_k^2}{p_k^2 - \omega_i^2} V_i(a_k) \tag{7}$$

$$A_{kj} = \frac{p_k^2}{p_k^2 - \omega_j^2} V_j(a_k) \tag{8}$$

Перемножая их левые и правые части, получим

$$A_{ki} A_{kj} = \frac{p_k^2}{p_k^2 - \omega_i^2} \frac{p_k^2}{p_k^2 - \omega_j^2} V_i(a_k) V_j(a_k) \tag{9}$$

Проинтегрируем вторые уравнения в (5) и (6) по длине стержня. Тогда эти уравнения приобретут вид

$$\begin{aligned} -\omega_i^2 \int_0^l V_i(x) V_j(x) dx + b \int_0^l \frac{d^2 V_i(x)}{dx^2} \frac{d^2 V_j(x)}{dx^2} dx = \\ = \sum_{k=1}^n e_k (A_{ki} - V_i(a_k)) V_j(a_k) \end{aligned} \tag{10}$$

$$\begin{aligned} & -\omega_j^2 \int_0^\ell V_j(x)V_i(x)dx + b \int_0^\ell \frac{d^2V_j(x)}{dx^2} \frac{d^2V_i(x)}{dx^2} dx = \\ & = \sum_{k=1}^n e_k (A_{kj} - V_j(a_k))V_i(a_k) \end{aligned} \quad (11)$$

Подставим (7), (8) в правые части (10) и (11)

$$\begin{aligned} & -\omega_i^2 \int_0^\ell V_i(x)V_j(x)dx + b \int_0^\ell \frac{d^2V_i(x)}{dx^2} \frac{d^2V_j(x)}{dx^2} dx = \\ & = \sum_{k=1}^n e_k \frac{P_k^2}{P_k^2 - \omega_i^2} V_i(a_k)V_j(a_k) - e_k V_i(a_k)V_j(a_k) \end{aligned} \quad (12)$$

$$\begin{aligned} & -\omega_j^2 \int_0^\ell V_j(x)V_i(x)dx + b \int_0^\ell \frac{d^2V_j(x)}{dx^2} \frac{d^2V_i(x)}{dx^2} dx = \\ & = \sum_{k=1}^n e_k \frac{P_k^2}{P_k^2 - \omega_j^2} V_j(a_k)V_i(a_k) - e_k V_j(a_k)V_i(a_k) \end{aligned} \quad (13)$$

Умножим на (-1) уравнение (12) и сложим с уравнением (13)

$$\begin{aligned} & (\omega_i^2 - \omega_j^2) \int_0^\ell V_i(x)V_j(x)dx = \sum_{k=1}^n e_k \left( \frac{-P_k^2}{P_k^2 - \omega_i^2} + \frac{P_k^2}{P_k^2 - \omega_j^2} \right) V_i(a_k)V_j(a_k) = \\ & = -\sum_{k=1}^n e_k \frac{(\omega_i^2 - \omega_j^2)}{(P_k^2 - \omega_i^2)(P_k^2 - \omega_j^2)} V_i(a_k)V_j(a_k) \end{aligned}$$

Переносим правую часть влево и вынося за скобки  $(\omega_i^2 - \omega_j^2)$ , получим

$$(\omega_i^2 - \omega_j^2) \left( \int_0^\ell V_i(x)V_j(x)dx + \sum_{k=1}^n e_k \frac{P_k^2}{(P_k^2 - \omega_i^2)(P_k^2 - \omega_j^2)} V_i(a_k)V_j(a_k) \right) = 0$$

Отсюда при  $i \neq j$  получим, что

$$\int_0^\ell V_i(x)V_j(x)dx + \sum_{k=1}^n e_k \frac{P_k^2}{(P_k^2 - \omega_i^2)(P_k^2 - \omega_j^2)} V_i(a_k)V_j(a_k) = 0 \text{ или, используя (9)}$$

$$\int_0^\ell V_i(x)V_j(x)dx + \sum_{k=1}^n \frac{e_k}{P_k} A_{ki}A_{kj} = 0 \quad (14)$$

Таким образом, условие ортогональности имеет вид

$$\int_0^\ell V_i(x)V_j(x)dx + \sum_{k=1}^n \frac{e_k}{P_k} A_{ki}A_{kj} = \begin{cases} 0, & i \neq j \\ \int_0^\ell V_i^2(x)dx + \sum_{k=1}^n \frac{e_k}{P_k} A_{ki}^2, & i = j \end{cases} \quad (15)$$

## 2. Решение начально-краевой задачи

Решение системы (1) ищем в виде разложений в ряды Фурье

$$z_k = \sum_{i=1}^{\infty} \varphi_i(t) A_{ki}, u(x, t) = \sum_{i=1}^{\infty} \varphi_i(t) V_i(x) \tag{16}$$

Подставляя в (1), получим

$$\begin{aligned} \ddot{\varphi}_i(t) A_{ki} + p_k^2 (\varphi_i(t) A_{ki} - \varphi_i(t) V_i(a_k)) &= 0 \\ \ddot{\varphi}_i(t) V_i(x) + b \varphi_i(t) \frac{d^4}{dx^4} V_i(x) &= \\ = \sum_{k=1}^N e_k (\varphi_i(t) A_{ki} - \varphi_i(t) V_i(x)) \delta(x - a_k) \end{aligned} \tag{17}$$

Разделив обе части двух уравнений в (3) на  $\varphi_i(t)$ , получим

$$\begin{aligned} \frac{\ddot{\varphi}_i(t)}{\varphi_i(t)} A_{ki} + p_k^2 (A_{ki} - V_i(a_k)) &= 0 \\ \frac{\ddot{\varphi}_i(t)}{\varphi_i(t)} V_i(x) + b \frac{d^4}{dx^4} V_i(x) &= \sum_{k=1}^N e_k (A_{ki} - V_i(x)) \delta(x - a_k) \end{aligned} \tag{18}$$

Обозначив  $\frac{\ddot{\varphi}_i(t)}{\varphi_i(t)} = -\omega_i^2$ , (19)

получим

$$\begin{aligned} -\omega_i^2 A_{ki} + p_k^2 (A_{ki} - V_i(a_k)) &= 0 \\ -\omega_i^2 V_i(x) + b \frac{d^4}{dx^4} V_i(x) &= \sum_{k=1}^N e_k (A_{ki} - V_i(x)) \delta(x - a_k) \end{aligned} \tag{20}$$

Таким образом парциальное движение является синусоидальным с частотой  $\omega_i$  со смещением

$$\varphi_i(t) = A_i \sin \omega_i t + B_i \cos \omega_i t, \tag{21}$$

которое получается в результате решения дифференциального уравнения  $\ddot{\varphi}_i(t) + \omega_i^2 \varphi_i(t) = 0$ , вытекающим из (19).

Запишем начальное условие

$$\begin{aligned} z_k(0) = \sum_{i=1}^{\infty} \varphi_i(0) A_{ki} = z_{ko}, u(x, 0) = \sum_{i=1}^{\infty} \varphi_i(0) V_i(x) = f_1(x) \\ \frac{dz_k}{dt}(0) = \sum_{i=1}^{\infty} \dot{\varphi}_i(0) A_{ki} = z_{tko}, \frac{\partial u}{\partial t}(0) = \sum_{i=1}^{\infty} \dot{\varphi}_i(0) V_i(x) = f_2(x) \text{ или} \\ \sum_{i=1}^{\infty} \varphi_i(0) A_{ki} = z_{ko} \end{aligned} \tag{22}$$

$$\sum_{i=1}^{\infty} \varphi_i(0) V_i(x) = f_1(x) \tag{23}$$

$$\sum_{i=1}^{\infty} \dot{\varphi}_i(0) A_{ki} = z_{tko} \tag{24}$$

$$\sum_{i=1}^{\infty} \dot{\varphi}_i(0) V_i(x) = f_2(x) \tag{25}$$

Умножим справа (22) на  $\frac{e_k A_{kj}}{p_k^2}$  и просуммируем по  $k$

$$\sum_{i=1}^{\infty} \varphi_i(0) \sum_{k=1}^N \frac{e_k A_{ki} A_{kj}}{p_k^2} = \sum_{k=1}^N \frac{e_k z_{ko} A_{kj}}{p_k^2} \tag{26}$$



Умножим (23) на  $V_j(x)$  и проинтегрируем по длине стержня

$$\sum_{i=1}^{\infty} \varphi_i(0) \int_0^{\ell} V_i(x) V_j(x) dx = \int_0^{\ell} f_1(x) V_j(x) dx \quad (27)$$

Сложив левые и правые части уравнений (26) и (27), получим

$$\begin{aligned} \sum_{i=1}^{\infty} \varphi_i(0) \left( \int_0^{\ell} V_i(x) V_j(x) dx + \sum_{k=1}^N \frac{e_k A_{ki} A_{kj}}{P_k^2} \right) = \\ = \int_0^{\ell} f_1(x) V_j(x) dx + \sum_{k=1}^N \frac{e_k z_{ko} A_{kj}}{P_k^2} \end{aligned} \quad (28)$$

В силу условия ортогональности (15) в левой части останется только слагаемое при  $i = j$ .

В результате получим

$$\varphi_i(0) = \frac{\int_0^{\ell} f_1(x) V_i(x) dx + \sum_{k=1}^N \frac{e_k z_{ko} A_{ki}}{P_k^2}}{\int_0^{\ell} V_i^2(x) dx + \sum_{k=1}^N \frac{e_k A_{ki} A_{ki}}{P_k^2}} \quad (29)$$

Аналогично для коэффициентов  $\varphi_{ti}(0)$  получим

$$\varphi_{ti}(0) = \frac{\int_0^{\ell} f_2(x) V_i(x) dx + \sum_{k=1}^N \frac{e_k z_{tko} A_{ki}}{P_k^2}}{\int_0^{\ell} V_i^2(x) dx + \sum_{k=1}^N \frac{e_k A_{ki} A_{ki}}{P_k^2}} \quad (30)$$

Из начальных условий  $\varphi_i(0) = \varphi_{i0}$ ,  $\varphi_{ti}(0) = \varphi_{ti0}$  и (21), получим

$$A_i = \frac{\varphi_{ti0}}{\omega_i}, \quad B_i = \varphi_{i0}.$$

В результате  $\varphi_i(t) = \frac{\varphi_{ti0}}{\omega_i} \sin \omega_i t + \varphi_{i0} \cos \omega_i t$  (31)

Подставляя в (16), получаем решение системы (1).

**Примечание.**  $V_i(x) = \sum_{k=1}^n \bar{V}_{ki}(x) A_{ki}$ , где  $\bar{V}_{ki}(x) = \frac{e_k \omega^2}{\omega^2 - P_k^2} \bar{V}_{ki}(x - a_k)$ , а  $\bar{V}_{ki}(x)$  - решения краевых задач

$$\begin{aligned} -\omega^2 \bar{V}(x) + b \frac{d^4 \bar{V}(x)}{dx^4} &= \delta(x), \\ \bar{V}(-a_k) &= \bar{V}(l - a_k) = 0, \\ \frac{d\bar{V}}{dx}(-a_k) &= \frac{d\bar{V}}{dx}(l - a_k) = 0, \\ k &= 1, 2, \dots, n \end{aligned}$$

### 3. Заключение

Таким образом получено решение начально-краевой задачи для произвольного числа осцилляторов, закрепленных на стержне в разных точках. По заданным начальным положениям и скоростям тел в осцилляторах, а также начальному изгибу стержня и скоростям точек стержня, изложенная методика позволяет определить полностью закон движения рассматриваемой механической системы, то есть найти в произвольный момент времени положение и скорости стержня и закрепленных на нем тел в осцилляторах.

### Литература

1. Баргуев С.Г., Мижидон А.Д., Цыцыренова М.Ж. Определение собственных частот простейшей механической системы на упругом основании. Вестник БГУ. Выпуск 9. Математика и информатика г.Улан-Удэ, 2009.

*Баргуев Сергей Ганжурович (Гаврилович)*, кандидат физико-математических наук, доцент Бурятского филиала Сибирского университета телекоммуникаций и информатики. E-mail: barguev@yandex.ru  
*Аюшеев Тумэн Владимирович*, доктор технических наук, доцент Восточно-Сибирского университета технологии и управления.

*Мижидон Арсалан Дугарович*, доктор технических наук, профессор Восточно-Сибирского университета технологии и управления. E-mail: miarsdu@esstu.ru

*Barguev Sergey Ganshurovich (Gavrilovich)*, candidate of physical and mathematical sciences, associate professor of Buryat branch of Siberian University telecommunication and information

*Ausheev Tumen Vladimirovich*, doctor of technical science, associate professor of East-Siberian State University of technology and management

*Mizhidon Arsalan Dugarovich*, doctor of technical science, professor of East-Siberian State University of technology and management.

## ОБ ОДНОЙ ЗАДАЧЕ НАКЛОННОЙ ПРОИЗВОДНОЙ С ЛИНЕЙНЫМИ КОЭФФИЦИЕНТАМИ

В данной статье доказана разрешимость одной из задач о наклонной производной для регулярных в единичном шаре гармонических функций.

**Ключевые слова:** гармоническая функция, оператор Лапласа, однородные полиномы.

*V. V. Kibirev*

### ON ONE PROBLEM OF THE OBLIQUE DERIVATIVE WITH LINEAR COEFFICIENTS

In this article the solvability of one of the problems of the oblique derivative for regular harmonic functions in the single sphere has been proved.

**Keywords:** harmonic functions, Laplace operator, homogeneous polynomials.

#### Введение

Задача о наклонной производной заключается в определении гармонической в области  $D$  функции  $U(x, y, z)$ , непрерывной вместе со своими производными первого порядка вплоть до границы  $S$  и удовлетворяющая условию  $(\text{grad } u \cdot P)_s = f$ , где  $f$  – заданная на  $S$  непрерывная функция

#### 1. Постановка задачи

В данной работе ставится такая задача: найти регулярную в шаре  $\Sigma: \{x^2 + y^2 + z^2 < 1\}$ , непрерывно дифференцируемую в замкнутом шаре  $\bar{\Sigma}$ , гармоническую функцию  $u$ , удовлетворяющую на сфере  $S: \{x^2 + y^2 + z^2 = 1\}$  условию

$$\lambda z u_z + (x + ay)u_x + (-ax + y)u_y = f, \quad (1)$$

где коэффициенты и функция  $f$  – непрерывно дифференцируемые на  $S$  функции.

#### 2. Теорема

Задача (1) при  $\lambda < 0$  имеет единственное решение, принимающее наперед заданные значения на многообразии точек выхода векторного поля  $(x + ay, -ax + y, \lambda z)$  в касательную к сфере плоскость.

**Доказательство:** Условие (1) перепишем так:

$$\lambda z u_z + x u_x + a y u_x - a x u_y + y u_y = f$$

$$\text{или } x u_x + y u_y + \lambda z u_z + a(y u_x - x u_y) = f \quad (1^*)$$

Как известно, что если функция  $u$  гармоническая, то функция  $W = x u_x + y u_y + \lambda z u_z + a(y u_x - x u_y)$  является бигармонической и имеет место представление  $W = g + (1 - x^2 - y^2 - z^2)h$ , где  $g$  и  $h$  – регулярные в шаре  $\Sigma$  гармонические функции [1]. Функции  $u, g, h$  связаны соотношениями:

$$\left. \begin{aligned} x u_x + y u_y + \lambda z u_z + a(y u_x - x u_y) &= g + (1 - r^2)h \\ (\lambda - 1)u_{zz} &= -2r \frac{\partial h}{\partial r} - 3h \end{aligned} \right\} \quad (2)$$

где  $r^2 = x^2 + y^2 + z^2$ . Второе уравнение последней системы получается, если к уравнению (1<sup>\*</sup>) применить оператор Лапласа. На сфере  $S$  функция  $g$  совпадает с  $f$ .

Можно показать, что для любой наперед заданной гармонической функции  $g$  всегда найдется гармоническая функция  $h$  такая, что система (2) будет справедлива для некоторой гармонической функции  $u$  [4].

Всякую регулярную в шаре гармоническую функцию  $H(x, y, z)$  можно представить абсолютно сходящимся в шаре рядом по шаровым функциям. Перегруппировав в этом ряде члены для  $H$  получим представление

$$H(x, y, z) = \sum_{l=0}^{\infty} H_l(x^2 + y^2, z) P_l(x, y) \tag{3}$$

где  $P_l$  – есть однородные гармонические полиномы степени  $l$  от двух переменных, а функция  $H_l(s, z)$  удовлетворяет уравнению

$$\frac{\partial^2 H_l}{\partial z^2} + 4S \frac{\partial^2 H_l}{\partial s^2} + 4(l+1) \frac{\partial H_l}{\partial s} = 0 \tag{4}$$

Для регулярной в шаре  $\Sigma$  гармонической функции  $H$  все функции  $H_l$  определяются однозначно аналитической в круге  $|z| < 1$  функцией  $H_l(0, z)$  одного комплексного переменного  $z$ .

Пусть  $u = \sum_{m=0}^{\infty} u_m$ ,  $g = \sum_{m=0}^{\infty} g_m$ ,  $h = \sum_{m=0}^{\infty} h_m$ , где

$$\left. \begin{aligned} u_m &= v_m(x^2 + y^2, z) \rho^m \cos m\alpha + W_m(x^2 + y^2, z) \rho^m \sin m\alpha \\ g_m &= p_m(x^2 + y^2, z) \rho^m \cos m\alpha + Q_m(x^2 + y^2, z) \rho^m \sin m\alpha \\ h_m &= \phi_m(x^2 + y^2, z) \rho^m \cos m\alpha + \Psi_m(x^2 + y^2, z) \rho^m \sin m\alpha \end{aligned} \right\} \tag{5}$$

Тогда для осесимметричных функций, аналогично тому, как и в [3] имеем:

$$\left. \begin{aligned} \lambda z \tilde{u}_z &= \tilde{g} + (1 - z^2) \tilde{h} \\ (\lambda - 1) \tilde{u}_{zz} &= -2z \frac{\partial \tilde{h}}{\partial z} - 3\tilde{h} \end{aligned} \right\} \tag{6}$$

где  $\tilde{u} = u(0, 0, z)$ ,  $\tilde{g} = g(0, 0, z)$ ,  $\tilde{h} = h(0, 0, z)$

Решая систему (6), получим, что в этом случае по известной функции  $\tilde{g}$  всегда можно найти функцию  $\tilde{h}$ .

Аналогично в общем случае можно получить и функцию  $u$ , в итоге получаем следующую теорему:

Задача (1) при  $\lambda < 0$  имеет единственное решение, принимающее наперед заданные значения на многообразии точек выхода векторного поля  $(x + ay, -ax + y, \lambda z)$  в касательную к сфере плоскость.

Эта задача может применяться в моделировании природных процессов, в частности, в теории приливов.

### Заключение

Таким образом, в данной работе доказана безусловная разрешимость задачи о наклонной производной для регулярных в единичном шаре гармонических функций с граничным условием (1).

### Литература

1. Бицадзе А.В. Краевые задачи для эллиптических уравнений второго порядка. М., Наука, 1966.
2. Коддингтон Э.А., Левинсон Н. Теория обыкновенных дифференциальных уравнений. Изд-во иностранной литературы, М., 1958.
3. Кибирев В.В. К задаче о наклонной производной с линейными коэффициентами для гармоничных функций // Дифференц. уравнения. 1980. Т.16, №1. С. 80-85
4. Янушаускас А. Применение вырождающихся уравнений к изучению задачи о наклонной производной // Дифференц. уравнения. 1969. Т.5, №1. С. 80-91.

*Кибирев Владимир Васильевич*, канд. физ.-мат. наук, проф. каф. прикладной математики Бурятского государственного университета, тел. (8301-2) 217573, [dekanat\\_imi@bsu.ru](mailto:dekanat_imi@bsu.ru)

*Kibirev Vladimir Vasilievich*, candidate of physical and mathematical sciences, the professor of the applied mathematics department of the Buryat State University.

## РАЗРЕШИМОСТЬ ОБРАТНОЙ ЗАДАЧИ ДЛЯ ПАРАБОЛИЧЕСКОГО УРАВНЕНИЯ ВЫСОКОГО ПОРЯДКА С НЕИЗВЕСТНЫМ КОЭФФИЦИЕНТОМ ПОГЛОЩЕНИЯ

Статья посвящена исследованию параболического уравнения высокого порядка  $u_t + u_{xxxx} + q(x)u = f(x, t)$ , где вместе с решением  $u(x, t)$  ищется также и коэффициент  $q(x)$ . При выполнении естественных граничных условий, некоторых условий переопределения, условий принадлежности входных данных определенному функциональному пространству доказывается теорема существования решения. При доказательстве используется комбинация методов линеаризации, продолжения по параметру и теорема Шаудера.

**Ключевые слова:** параболические уравнения высокого порядка, обратные задачи математической физики.

O. Yu. Nikolaev

## RESOLVABILITY OF A RETURN TASK FOR THE PARABOLIC EQUATION OF A HIGH ORDER WITH UNKNOWN FACTOR OF ABSORPTION

Article is devoted to research of the parabolic equation of a high order  $u_t + u_{xxxx} + q(x)u = f(x, t)$ , where together with the decision  $u(x, t)$  is looked for as well factor of  $q(x)$ . At performance of the natural boundary conditions, some conditions of redefinition, conditions of accessory of entrance data a certain functional space proves the theorem of existence of the decision. At the proof the combination of methods of linearization, continuation on parameter and Shauder's theorem is used.

**Keywords:** parabolic equations of a high order, return problems of mathematical physics.

### Введение

Исследуется задача нахождения вместе с решением  $u(x, t)$  параболического уравнения высокого порядка

$$u_t + u_{xxxx} + q(x)u = f(x, t)$$

также коэффициента  $q(x)$ . При выполнении естественных граничных условий, начального условия и условия переопределения, принадлежности входных данных определенному функциональному пространству доказывается теорема существования решения. Ранее подобные задачи изучались при специальных (менее общих, чем в настоящей работе) условиях, в статьях [1, 2].

### 1. Постановка задачи.

Пусть область  $Q$  есть прямоугольник  $\{(x, t) | x \in (0, 1), t \in (0, T)\}$ ,  $\Omega = \{x | x \in (0, 1)\}$ . Функции  $f(x, t)$ ,  $u_0(x)$ ,  $u_1(x)$ ,  $\varphi_0(t)$ ,  $\varphi_1(t)$ ,  $\varphi_2(t)$ ,  $\varphi_3(t)$  заданы при  $f(x, t) \in L_2(Q)$ ,  $f_t(x, t) \in L_2(Q)$ .

**Обратная задача:** найти функции  $u(x, t)$ ,  $q(x)$ , связанные в прямоугольнике  $Q$  уравнением

$$u_t + u_{xxxx} + q(x)u = f(x, t), \quad (1)$$

при выполнении для функции  $u(x, t)$  условий

$$u(x, 0) = u_0(x), x \in \Omega, \quad (2)$$

$$u(x, T) = u_1(x), x \in \Omega, \quad (3)$$

$$u(0, t) = \varphi_0(t), u(1, t) = \varphi_1(t), t \in (0, T) \quad (4)$$

$$u_x(0, t) = \varphi_2(t), u_x(1, t) = \varphi_3(t), t \in (0, T). \quad (5)$$

### 2. Разрешимость обратной задачи.

Определим пространства  $V, V_1$ :

$$V = \{v(x, t) : v(x, t) \in L_2(Q), v_t(x, t) \in L_2(Q), v_{xxxx} \in L_2(Q),$$

$$v(x, T) \in L_2(\Omega), v(0, t) = v(1, t) = v_x(0, t) = v_x(1, t), t \in (0, T)\}, \text{ с нормами}$$

$$V_1 = \{v(x, t) : v(x, t) \in V, v_t(x, t) \in V\}$$

$$\|v\|_V = \left[ \int_Q (v^2 + v_t^2 + v_{xxxx}^2) dxdt + \int_{\Omega} v_{xx}^2(x, T) dx \right]^{\frac{1}{2}},$$

$$\|v\|_{V_1} = \|v\|_V + \|v_t\|_V.$$

С помощью сдвига

$$v(x, t) = u(x, t) - U(x, t),$$

$$U(x, t) = -2\varphi_1(t) + \varphi_3(t) + 2\varphi_0(t) x^3 +$$

$$+ 3\varphi_1(t) - \varphi_3 t - 2\varphi_2 t - 3\varphi_0 t x^2 + \varphi_2 t x + \varphi_0 t,$$

уравнение (1) и начально-краевые условия (2), (3), (4), (5) преобразуются к виду

$$v_t + v_{xxxx} + q(x)(v + U) = f_1(x, t), \tag{6}$$

$$f_1(x, t) = f(x, t) - U_t(x, t),$$

$$v(x, 0) = v_0(x), v_0(x) = u_0(x) - U(x, 0), x \in \Omega, \tag{7}$$

$$v(x, T) = v_1(x), v_1(x) = u_1(x) - U(x, T), x \in \Omega, \tag{8}$$

$$v(0, t) = v(1, t) = u_x(0, t) = v_x(1, t) = 0, t \in [0, T]. \tag{9}$$

Положим  $\alpha(x) = \frac{f_1(x, T) - u_{1xxxx}(x)}{u_1(x)}, \beta(x) = -\frac{1}{u_1(x)},$

$$a(x) = \frac{u_0(x)}{u_1(x)}, b(x) = f_1(x, 0) - u_{0xxxx}(x) - \frac{u_0(x)}{u_1(x)} [f_1(x, T) - u_{1xxxx}(x)].$$

Пусть числа  $k_0, k_1, k_2, k_4, \alpha_0, \beta, \mu_0$  - некоторые положительные числа, роль которых мы проясним ниже. Определим также числа  $\bar{\alpha}, \bar{\beta}, \bar{a}, \bar{a}_1, \bar{a}_2, \bar{b}, \bar{b}_2, \bar{U}_1, \bar{f}_1, c_0, B, K_6$ :

$$\bar{\alpha} = \|\alpha(x)\|_{L_{\infty}(\Omega)}, \bar{\beta} = \|\beta(x)\|_{L_{\infty}(\Omega)}, \bar{U}_1 = \|U_t(x, t)\|_{L_{\infty}(Q)}, \bar{a} = \|a(x)\|_{L_{\infty}(\Omega)},$$

$$\bar{b} = \|b(x)\|_{L_{\infty}(\Omega)}, \bar{a}_1 = \|a_x(x)\|_{L_{\infty}(\Omega)}, \bar{a}_2 = \|a_{xx}(x)\|_{L_{\infty}(\Omega)}, \bar{b}_2 = \|b_{xx}(x)\|_{L_{\infty}(\Omega)},$$

$$\bar{f}_1 = \|f_{1t}(x, t)\|_{L_2(Q)}, c_0 = \bar{a}\bar{a}_2 + \bar{a}\bar{b}_2 + 2\bar{a}_1^2 + a_1 a_2 + 2\bar{a}_1 \bar{b}_2,$$

$$B = \frac{12^3 \bar{a} \bar{a}_1^{-4}}{k_1^3} + \frac{4c_0}{k_1} \frac{\bar{a}\bar{a}_2 + 2\bar{a}_1^2 + 2\bar{a}_1 \bar{b}_2}{\bar{a}\bar{a}_2 + 2\bar{a}_1^2 + 2\bar{a}_1 \bar{b}_2} + \bar{a}_2 \bar{b}_2 + \bar{a}_1 \bar{a}_2 \left( \frac{k_1}{4c_0} + \frac{8c_0}{k_1} \right),$$

$$K_6 = \frac{1}{k_4} \left[ \frac{1}{k_2} B + 4 \bar{\alpha} + \alpha_0^2 + 4 \right] \bar{\beta}^2 \bar{U}_1 T.$$

**Теорема 1.** Пусть выполняются условия

$$u_0(x) \in C^4(\bar{Q}), u_1(x) \in C^4(\bar{Q}), \varphi_i(t) \in C[0, T], i = 1, \dots, 4,$$

$$K_6 < 1 \tag{10}$$

$$|u_1(x)| \geq k_0 > 0, x \in \bar{Q}, \tag{11}$$

$$1 - \bar{a}^2 \geq k_1 > 0, \tag{12}$$

$$|\alpha(x)| \geq \alpha_0 > 0, |\beta(x)| \leq \beta, x \in \bar{Q}, \tag{13}$$

$$0 < \mu_0 \leq \frac{\alpha_0}{\beta}. \tag{14}$$

Тогда для любой функции  $f(x, t)$  такой, что  $f(x, t) \in L_2(Q), f_t(x, t) \in L_2(Q)$ , обратная задача имеет решение  $u(x, t) \in W_2^{4,1}(Q), q(x) \in L_2(\Omega)$ .

**Доказательство.** Перейдем к прямой задаче нахождения  $v(x, t)$ . Из уравнения (1) при  $t=T$  имеем представление для  $q(x)$ :

$$q(x) = \frac{f_1(x,t) - u_{xxxx} x - v_t x, T}{u_1(x)}, q(x) = \alpha(x) + \beta(x)v_t(x, T).$$

При  $t=0$  выполняется

$$v_t(x, 0) = a(x)v_t(x, T) + b(x). \quad (15)$$

Продифференцируем уравнение (6) по  $t$ :

$$v_{tt} + v_{xxxx} + \alpha(x) + \beta(x)v_t(x, T) \quad v_t + V_t = f_{1t}(x, t). \quad (16)$$

Рассмотрим краевую задачу: найти функцию  $v(x, t)$ , являющуюся в прямоугольнике  $Q$  решением уравнения (16) и удовлетворяющую условиям (7), (8), (9), (15). Разрешимость данной краевой задачи докажем комбинируя метод срезающих функций, метод продолжения по параметру и метод неподвижной точки. Именно с помощью решения данной краевой задачи и будет построено решение исходной обратной задачи.

Определим срезающую функцию  $G(\xi)$ :

$$G(\xi) = \begin{cases} \xi, & \text{если } |\xi| \leq \mu_0 \\ \mu_0, & \text{если } \xi > \mu_0; \\ -\mu_0, & \text{если } \xi < -\mu_0. \end{cases}$$

Применим операцию срезки к уравнению (16) и рассмотрим краевую задачу: найти функцию  $v(x, t)$ , являющуюся в прямоугольнике  $Q$  решением уравнения

$$\begin{aligned} v_{tt} + v_{xxxx} + [\alpha(x) + \beta(x)G(v_t, x, T)]v_t = \\ = f_{1t}(x, t) - [\alpha(x) + \beta(x)G(v_t, x, T)]U_t \end{aligned} \quad (17)$$

и удовлетворяющую условиям (7), (8), (9), (15). Разрешимость этой задачи мы докажем с помощью метода продолжения по параметру и теоремы Шаудера.

Пусть  $z(x, t)$  есть произвольная функция из пространства  $V_1$ ,  $\lambda$  есть число из отрезка  $[0, 1]$ . Рассмотрим еще одну краевую задачу: найти функцию  $v(x, t)$ , являющуюся в прямоугольнике  $Q$  решением уравнения

$$\begin{aligned} v_{tt} + v_{xxxx} + [\alpha(x) + \beta(x)G(z_t, x, T)]v_t = \\ = f_{1t}(x, t) - [\alpha(x) + \beta(x)G(z_t, x, T)]U_t \end{aligned} \quad (17_z)$$

и такую, что для нее выполняются условия (7), (8), (9), а также следующее условие:

$$v_t(x, 0) = \lambda a(x)v_t(x, T) + b(x). \quad (15_\lambda)$$

При  $\lambda=0$  данная задача есть начально-краевая задача для линейного параболического относительно функции  $v_t(x, t)$  уравнения, разрешимость ее в пространстве  $V$  при выполнении некоторых условий согласования и гладкости (которые будут указаны ниже) известна, см. [3]. Зная функцию  $v_t(x, t)$ , нетрудно найти саму функцию  $v(x, t)$  из пространства  $V_1$ , являющуюся решением задачи (17<sub>z</sub>) с условиями (7), (8), (9), (15<sub>λ</sub>) при  $\lambda=0$ . Для того, чтобы получить разрешимость задачи (17<sub>z</sub>), (7), (8), (9), (15<sub>λ</sub>) при всех  $\lambda$  (в том числе и при  $\lambda = 1$ ), достаточно установить наличие равномерной по  $\lambda$  априорной оценки возможных решений этой задачи в пространстве  $V_1$ .

Обозначим  $\omega = v_t$ .

Для получения первой априорной оценки умножим уравнение (17<sub>z</sub>) на  $\omega$  и проинтегрируем обе части по области  $Q$ . Интегрируя по частям, получим следующее равенство:

$$\begin{aligned} & \frac{1}{2} \int_{\Omega} \omega^2(x, T) dx - \frac{1}{2} \int_{\Omega} \omega^2(x, 0) dx + \int_0^T \int_{\Omega} \omega_{xx}^2 dx d\tau + \\ & + \int_0^T \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] \omega^2 dx d\tau = \\ & = \int_0^T \int_{\Omega} f_{1\tau}(x, \tau) \omega dx d\tau - \int_0^T \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] U_{\tau} \omega dx d\tau. \end{aligned}$$

В силу (13) и (14) выполняется неравенство

$$\alpha(x) + \beta(x)G(z_t(x, T)) \geq 0, x \in 0, 1,$$

что позволяет оценить интеграл

$$\int_0^T \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] \omega^2 dx d\tau \geq 0.$$

Используя неравенство Юнга и условие (15<sub>1</sub>), получаем оценки

$$\begin{aligned} \int_{\Omega} \omega^2(x, 0) dx &\leq \bar{a}^2 (1 + \delta_1^2) \int_{\Omega} \omega^2(x, T) dx + (1 + \frac{1}{\delta_1^2}) \bar{b}^2, \\ \left| \int_0^T \int_{\Omega} [\alpha(x) + \beta(x)G(z_t)] U_t \omega dx d\tau \right| &\leq (\frac{\delta_2^2}{2} + \frac{\delta_3^2}{2}) \int_0^T \int_{\Omega} \omega^2 dx d\tau + \\ &+ \frac{\bar{\alpha}^2 \bar{U}_1^2}{2\delta_2^2} + \frac{\bar{\beta}^2 U_1^2}{2\delta_3^2} T \|z_t(x, T)\|_{L_x(\Omega)}^2, \\ \left| \int_0^T \int_{\Omega} f_{1\tau}(x, \tau) \omega(x, \tau) dx d\tau \right| &\leq \frac{\delta_4^2}{2} \int_0^T \int_{\Omega} \omega^2(x, \tau) dx d\tau + \frac{\bar{f}_1^2}{2\delta_4^2}. \end{aligned}$$

Можно оценить

$$\begin{aligned} \frac{1}{2} \int_{\Omega} \omega^2(x, T) dx + \int_0^T \int_{\Omega} \omega_{xx}^2(x, \tau) dx d\tau &\leq \frac{1}{2} \bar{a}^2 (1 + \delta_1^2) \int_{\Omega} \omega^2(x, T) dx + \\ + \frac{\delta_2^2 + \delta_3^2 + \delta_4^2}{2} \int_0^T \int_{\Omega} \omega^2(x, \tau) dx d\tau &+ \frac{1}{2} (1 + \frac{1}{\delta_1^2}) \bar{b}^2 + \frac{\bar{\alpha}^2 \bar{U}_1^2}{2\delta_2^2} + \frac{\bar{f}_1^2}{2\delta_4^2} + \frac{\bar{\beta}^2 \bar{U}_1^2}{2\delta_3^2} T \|z_t(x, T)\|_{L_x(\Omega)}^2. \end{aligned}$$

Число  $\delta_1$  выберем с

учетом условия (12):  $\delta_1^2 = \frac{k_1}{2\bar{a}^2}.$

Положим  $\delta_2^2 = \delta_3^2 = \delta_4^2 = \frac{1}{2}, k_2 = \min(\frac{k_1}{4}, \frac{1}{4}).$

Первая априорная оценка будет иметь следующий вид:

$$\begin{aligned} \int_{\Omega} \omega^2(x, T) dx + \int_0^T \int_{\Omega} \omega_{xx}^2 dx d\tau &\leq K_1 + K_2 \|z_t(x, T)\|_{L_x(\Omega)}^2, \\ K_1 &= \frac{1}{k_2} (\bar{\alpha}^2 \bar{U}_1^2 + \bar{f}_1^2 + \frac{1}{2} (1 + \frac{2\bar{a}^2}{k_1}) \bar{b}^2), \\ K_2 &= \frac{1}{k_2} \bar{\beta}^2 \bar{U}_1^2 T. \end{aligned}$$

Для получения второй априорной оценки умножим уравнение (17<sub>z</sub>) на  $\omega_t$  и проинтегрируем обе части по области  $Q$ . Интегрируя по частям, получим следующее равенство:

$$\begin{aligned} \int_0^T \int_{\Omega} \omega_{\tau}^2 dx d\tau + \frac{1}{2} \int_{\Omega} \omega_{xx}^2(x, T) dx + \frac{1}{2} \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] \omega^2(x, T) dx &= \\ = \frac{1}{2} \int_{\Omega} \omega_{xx}^2(x, 0) dx + \frac{1}{2} \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] \omega^2(x, 0) dx + \\ + \int_0^T \int_{\Omega} f_{1\tau}(x, \tau) \omega dx d\tau - \int_0^T \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] U_{\tau} \omega dx d\tau. \end{aligned}$$

Применяя неравенство Юнга и условие (15<sub>1</sub>), придем к следующим неравенствам:



$$\begin{aligned}
 & \int_{\Omega} \left[ \alpha x + \beta x G z_t x, T \right] \omega^2 x, 0 dx = \\
 & = \int_{\Omega} \left[ \alpha x + \beta x G z_t x, T \right] a^2 x \omega^2 x, T + 2a x \omega x, T b a + b^2 x dx \leq \\
 & \leq \int_{\Omega} \left[ \alpha x + \beta x G z_t x, T \right] \left( 1 + \delta_0^2 a^2 x \omega^2 x, T + \left( 1 + \frac{1}{\delta_0^2} \right) b^2 x \right) dx, \\
 & \int_{\Omega} \omega_{xx}^2 x, 0 dx = \int_{\Omega} (\omega_{xx}^2 x, T a^2 x + 4\omega_{xx} x, T \omega_x x, T a x a_x x + \\
 & + 2\omega_{xx} x, T \omega x, T a x a_{xx} x + 2\omega_{xx} x, T a(x)b_{xx}(x) + 4\omega_x^2(x, T)a_x^2(x) + \\
 & + 4\omega_x(x, T)\omega(x, T)a_x(x)a_{xx}(x) + 4\omega_x(x, T)a_x(x)b_{xx}(x) + \\
 & + \omega^2(x, T)a_{xx}^2 + 2\omega(x, T)a_{xx}(x)b_{xx}(x) + b_{xx}^2(x)) \leq \\
 & \leq A \int_{\Omega} \omega_{xx}^2(x, T) dx + B_{\delta} \int_{\Omega} \omega^2(x, T) dx + C. \\
 & A = \bar{a}^{-2} + \bar{a}\bar{a}_1 \left( 2\delta_1^2 + \frac{\delta_2^2}{\delta_1^2} \right) + \bar{a}\bar{a}_2\delta_3^2 + \bar{a}\bar{b}_4^2 + 2\bar{a}_1^2\delta_5^2 + \bar{a}_1\bar{a}_2\delta_6^2\delta_7^2 + 2\bar{a}_1\bar{b}_2\delta_8^2, \\
 & B_{\delta} = \frac{\bar{a}\bar{a}_1}{\delta_1^2\delta_2^2} + \frac{\bar{a}\bar{a}_2}{\delta_3^2} + \frac{2\bar{a}_1^2}{\delta_5^2} + \bar{a}\bar{a}_2 \left( \frac{\delta_6^2}{\delta_7^2} + \frac{2}{\delta_6^2} \right) + \bar{a}_2^2 + \frac{2\bar{a}_1\bar{b}_2}{\delta_8^2} + \bar{a}_2\bar{b}_2\delta_9^2, \\
 & C = \frac{\bar{a}_2\bar{b}_2}{\delta_9^2} + \bar{b}_2^2 + \frac{\bar{a}\bar{b}_2}{\delta_4^2}.
 \end{aligned}$$

Следовательно,

$$\begin{aligned}
 & \left( 1 - \frac{\delta_{10}^2 + \delta_{11}^2 + \delta_{12}^2}{2} \right) \int_0^T \int_{\Omega} \omega_T^2 dx dt + \frac{1}{2} (1 - A) \int_{\Omega} \omega_{xx}^2 x, T dx + \\
 & + \frac{1}{2} \int_{\Omega} \left[ \alpha x + \beta x G z_t x, T \right] \omega^2 x, T dx \leq \\
 & \leq \frac{1}{2} \int_{\Omega} \left[ \alpha x + \beta x G z_t x, T \right] \left( 1 + \delta_0^2 a^2 x \omega^2 x, T + \left( 1 + \frac{1}{\delta_0^2} \right) b^2 x \right) dx + \\
 & + \frac{1}{2} B_{\delta} \int_{\Omega} \omega^2 x, T dx + A_1 + A_2 \|z_t x, T\|_{L_x \Omega}^2, \\
 & A_1 = \frac{1}{2\delta_{10}^2} \bar{f}_1^2 + \frac{1}{2\delta_{11}^2} \bar{\alpha} \bar{U}_1^2 + \frac{C}{2}, A_2 = \frac{1}{2\delta_{12}^2} \bar{\beta}^2 \bar{U}_1^2 T.
 \end{aligned}$$

Условие (12) позволяет положить

$$\begin{aligned}
 \delta_0 &= \frac{k_1}{2a}, \delta_1^2 = \delta_2 = \frac{k_1}{12aa_1}, \delta_3^2 = \delta_4^2 = \delta_5^2 = \delta_6^2 = \delta_8^2 = \frac{k_1}{4c_0}, \\
 \delta_7 &= \delta_9 = 1, \delta_{10}^2 = \delta_{11}^2 = \delta_{12}^2 = \frac{1}{2}.
 \end{aligned}$$

Тогда  $1 - A \geq \frac{k_1}{2}$ ,  $k_3 = \min\left(\frac{k_1}{4}, \frac{1}{4}\right)$ .

Вторая оценка будет иметь вид:

$$\begin{aligned}
 & \int_0^T \int_{\Omega} \omega_T^2 dx dt + \int_{\Omega} \omega_{xx}^2 x, T dx \leq K_3 + K_4 \|z_t x, T\|_{L_x \Omega}^2, \\
 & K_3 = \frac{1}{k_3} \left[ \frac{\bar{b}^2}{2} \left( 1 + \frac{2\bar{a}^2}{k_1} \right) \left( \bar{\alpha} + \frac{\bar{\beta}^2}{2} \right) + \frac{1}{2} K_1 B + \bar{f}_1^2 + \bar{\alpha} \bar{U}_1^2 + \frac{1}{2} \bar{a}_2 \bar{b}_2 + \frac{1}{2} \bar{b}_2^2 + \frac{2c_0 \bar{a} \bar{b}_2}{k_1} \right], \\
 & K_4 = \frac{1}{k_3} \left[ \frac{\bar{b}^2}{4} \left( 1 + \frac{2\bar{a}^2}{k_1} \right) + \frac{B \bar{\beta}^2 \bar{U}_1^2}{2k_2} T + \bar{\beta}^2 \bar{U}_1^2 T \right].
 \end{aligned}$$

Третья оценка для  $\omega(x, t)$  получается после умножения уравнения (17<sub>z</sub>) на  $\omega_{xxxx}(x, t)$  и интегрирования по области  $Q$ . Приходим к равенству:

$$\begin{aligned} & \int_0^T \int_{\Omega} \omega_{xxxx}^2 dx d\tau + \frac{1}{2} \int_{\Omega} \omega_{xx}^2(x, T) dx + \int_0^T \int_{\Omega} [\alpha x + \beta x G(z_t, x, T)] \omega \omega_{xxxx} dx d\tau = \\ & = \frac{1}{2} \int_{\Omega} \omega_{xx}^2(x, 0) dx + \int_0^T \int_{\Omega} f_{1\tau} \omega_{xxxx} dx d\tau - \int_0^T \int_{\Omega} [\alpha(x) + \beta(x)G(z_t(x, T))] U_{\tau} \omega_{xxxx} dx d\tau. \end{aligned}$$

Оценим сверху функцию  $\alpha x + \beta x G(z_t, x, T)$  :

$$\alpha x + \beta x G(z_t, x, T) \leq \bar{\alpha} + \bar{\beta} |z_t(x, T)| \leq \bar{\alpha} + \bar{\beta} \frac{\alpha_0}{\beta} = \bar{\alpha} + \alpha_0.$$

Применяя данную оценку, неравенство Юнга, первую и вторую оценки, а также представление для  $\int_{\Omega} \omega_{xx}^2(x, 0) dx$ , можно записать:

$$\begin{aligned} & \int_0^T \int_{\Omega} \omega_{xxxx}^2 dx d\tau + \frac{k_1}{4} \int_{\Omega} \omega_{xx}^2(x, T) dx \leq \\ & \leq \frac{1}{2} B \int_{\Omega} \omega^2(x, T) dx + \frac{1}{2} \bar{a}_2 \bar{b}_2 + \frac{1}{2} \bar{b}_2^2 + \frac{2c_0}{k_1} \bar{a} \bar{b}_2 + \\ & + \frac{\delta_1^2 + \delta_2^2 + \delta_3^2 + \delta_4^2}{2} \int_0^T \int_{\Omega} \omega_{xxxx}^2 dx d\tau + \frac{\bar{\alpha} + \alpha_0}{2\delta_1^2} \int_0^T \int_{\Omega} \omega^2 dx d\tau + \\ & + \frac{1}{\delta_2^2} \bar{f}_1 + \frac{1}{2\delta_3^2} \bar{\alpha} \bar{U}_1 + \frac{1}{2\delta_4^2} \bar{\beta} \bar{U}_1 T \|z_t(x, T)\|_{L_{\infty}(\Omega)}^2. \end{aligned}$$

Зафиксируем

$$\delta_1^2 = \delta_2^2 = \delta_3^2 = \delta_4^2 = \frac{1}{4}, k_4 = \min\left(\frac{1}{2}, \frac{k_1}{4}\right).$$

В результате получим третью априорную оценку:

$$\begin{aligned} & \int_0^T \int_{\Omega} \omega_{xxxx}^2 dx d\tau + \int_{\Omega} \omega_{xx}^2(x, T) dx \leq K_5 + K_6 \|z_t(x, T)\|_{L_{\infty}(\Omega)}^2, \\ & K_5 = \frac{1}{k_4} \left[ B + 4 \bar{\alpha} + \alpha_0^2 K_1 + \bar{a}_2 \bar{b}_2 + \bar{b}_2^2 + \frac{4c_0}{k_1} \bar{a} \bar{b}_2 + 4\bar{f}_1 + 4\bar{\alpha} \bar{U}_1 \right], \\ & K_6 = \frac{1}{k_4} \left[ \frac{1}{k_2} B + 4 \bar{\alpha} + \alpha_0^2 + 4 \right] \bar{\beta}^2 \bar{U}_1 T. \end{aligned}$$

Используя представление

$$v(x, t) = \int_0^t \omega(x, \tau) d\tau + v(x, 0),$$

нетрудно получить суммарную оценку  $v(x, t)$  в пространстве  $V_1$ :

$$\|v(x, t)\|_{V_1}^2 \leq N_1 + N_2 \|z_t(x, T)\|_{L_{\infty}(\Omega)}^2,$$

$$N_1 = 2 + 4T K_1 + K_3 + 1 + 2T K_5 + 2T \|v_0\|_{L_2(\Omega)}^2 + \|v_{0xx}\|_{L_2(\Omega)}^2 + \|v_{0xxxx}\|_{L_2(\Omega)}^2,$$

$$N_2 = 2 + 4T K_2 + K_4 + 1 + 2T K_6.$$

Т.к. выполняется неравенство

$$\|z_t(x, T)\|_{L_{\infty}(\Omega)}^2 \leq \|z(x, t)\|_{V_1}^2,$$

то краевая задача (17<sub>z</sub>), (7), (8), (9), (15<sub>λ</sub>) порождает оператор  $\Phi$ , действующий из пространства  $V_1$  в себя:

$$\Phi(z)=v.$$

Покажем, что этот оператор имеет в пространстве  $V_1$  неподвижные точки.

Пусть  $R_0, R_1$  есть некоторые положительные числа,  $W$  есть множество

$$W = \left\{ v \mid x, t \in V_1 : \|v_{xxt}(x, T)\|_{L_2 \Omega} \leq R_0, \left[ \int_Q v^2 + v_t^2 + v_{xxx}^2 + v_{tt}^2 + v_{xxxx}^2 dxdt \right]^{\frac{1}{2}} \leq R_1 \right\}.$$

Выполнение условий теоремы позволяет подобрать числа  $R_0, R_1$  так, что оператор  $\Phi$  будет переводить множество  $W$  в себя.

Выберем произвольный элемент  $z(x, t)$  из множества  $W$ . Тогда

$$\begin{aligned} \|z_{xxt}(x, T)\|_{L_2 \Omega}^2 &\leq R_0^2 \\ \|z_t(x, T)\|_{L_\infty \Omega}^2 &\leq \|z_{xxt}(x, T)\|_{L_2 \Omega}^2. \end{aligned}$$

Из третьей априорной оценки имеем

$$\int_\Omega v_{xxt}^2(x, T) dx \leq K_5 + K_6 \|z_t(x, T)\|_{L_\infty \Omega}^2.$$

При выполнении условия (10) будет выполняться неравенство

$$\begin{aligned} K_5 + K_6 \|z_t(x, T)\|_{L_\infty \Omega}^2 &\leq R_0^2 \\ R_0^2 &= \frac{K_5}{1 - K_6}. \end{aligned}$$

В качестве числа  $R_1$  возьмем произвольное число из промежутка  $[N_1 + N_2 R_0^2]^{\frac{1}{2}}, +\infty)$ . Тогда для решения краевой задачи (17<sub>z</sub>), (7), (8), (9), (15<sub>λ</sub>) будет выполняться неравенство

$$\int_Q v^2 + v_t^2 + v_{xxx}^2 + v_{tt}^2 + v_{xxxx}^2 dxdt \leq N_1 + N_2 R_0^2 \leq R_1^2$$

и оператор  $\Phi$  будет переводить множество  $W$  в себя.

Доказательство вполне непрерывности оператора  $\Phi$  в пространстве  $V_1$  при наличии полученных выше априорных оценок аналогично доказательству в работе [4]. Замкнутость, выпуклость и ограниченность множества  $W$  очевидна. Согласно теореме Шаудера, оператор  $\Phi$  имеет в множестве  $W$  неподвижную точку, т.е. существует функция  $v(x, t) \in W$ , являющаяся решением краевой задачи (17<sub>z</sub>), (7), (8), (9), (15<sub>λ</sub>). Наличие равномерных по  $\lambda$  априорных оценок решения краевой задачи (17<sub>z</sub>), (7), (8), (9), (15<sub>λ</sub>) является достаточным условием существования решения  $v(x, t)$  краевой задачи (17<sub>z</sub>), (7), (8), (9), (15).

Для найденной функции  $v(x, t)$  имеет место оценка

$$\|v_t(x, T)\|_{L_\infty \Omega} \leq \mu_0.$$

Поэтому для  $v(x, t)$  выполняется равенство

$$G v_t(x, T) = v_t(x, T).$$

Другими словами, решение  $v(x, t)$  краевой задачи (17), (7), (8), (9), (15) является решением краевой задачи (16), (7), (8), (9), (15).

Выполним интегрирование в уравнении (16) по временной переменной от 0 до текущей точки и определим функции  $u(x, t), q(x)$  следующим образом:

$$\begin{aligned} u(x, t) &= v(x, t) + U(x, t), \\ q(x) &= \alpha(x) + \beta(x)v_t(x, T). \end{aligned}$$

Очевидно, что функции  $u(x,t), q(x)$  связаны в области  $Q$  уравнением (1). Учитывая выполнение для функции  $u(x,t)$  условий (2), (3), (4), (5), а также полученную по ходу доказательства принадлежность функций  $v(x,t)$  и  $v_t(x,T)$  пространствам  $V_1$  и  $L_2(\Omega)$  соответственно, мы окончательно получим, что функции  $u(x,t), q(x)$  представляют собой требуемое решение обратной задачи. Теорема доказана.

#### Заключение

Доказана разрешимость обратной задачи для параболического уравнения высокого порядка с неизвестным коэффициентом поглощения при ненулевых начальных и краевых условиях с нелокальным условием переопределения.

#### Литература

1. Кириллова Г.А. Обратные задачи для параболических уравнений высокого порядка: автореф. дис. ... канд. физ.-мат. наук. – Рубцовск, 2004. – 102 с.
2. Кожанов А.И., Кириллова Г.А. О некоторых обратных задачах для параболического уравнения четвертого порядка // Мат. заметки ЯГУ. 2000. Т. 7, вып.1. С. 35-48.
3. Якубов С.Я. Линейные дифференциально-операторные уравнения и их приложения. Баку: Элм, 1985.
4. Кожанов А.И. Нелинейные нагруженные уравнения и обратные задачи // Журн. вычислит. математики и мат. физики. 2004. Т. 44, № 4. С. 694-716.

*Николаев Олег Юрьевич*, преподаватель кафедры математического анализа и методики преподавания математики ИМИ БГУ, тел. 8 914 6385317, e-mail: [nikolaev.oleg1@yandex.ru](mailto:nikolaev.oleg1@yandex.ru)

*Nikolaev Oleg Yurievich*, the teacher of chair of the mathematical analysis and a technique of teaching of mathematics IMI BSU.

## АЛГОРИТМ ИНТЕГРИРОВАНИЯ ПЕРЕМЕННОЙ КОНФИГУРАЦИИ НА ОСНОВЕ ЯВНО-НЕЯВНЫХ СХЕМ ЧЕТВЕРТОГО ПОРЯДКА<sup>1</sup>

Построены  $L$ -устойчивый (4,2)-метод и явная пятистадийная схема типа Рунге-Кутты четвертого порядка точности. Создан алгоритм интегрирования переменного шага, в котором выбор эффективной численной схемы осуществляется на каждом шаге с применением неравенства для контроля устойчивости. Приведены результаты расчетов, подтверждающие эффективность построенного алгоритма.

**Ключевые слова:** жесткие задачи, явный и неявный методы, контроль точности и устойчивости.

*A.E. Novikov, V.V. Shaidurov*

## AN INTEGRATION ALGORITHM OF VARIABLE CONFIGURATION BASED ON EXPLICIT-IMPLICIT SCHEMES OF 4<sup>TH</sup> ORDER OF ACCURACY

An  $L$ -stable (4,2)-method of 4th order of accuracy and an explicit Runge-Kutta scheme of 4th order of accuracy are constructed. An integration algorithm of variable step size is formulated. The most effective numerical scheme is chosen for each step by means of stability control. The numerical results which confirm the effectiveness of the algorithm are given.

**Key words:** stiff problems, explicit and implicit methods, stability and accuracy control

### Введение

Во многих важных приложениях возникает проблема численного решения задачи Коши для обыкновенных дифференциальных уравнений. В современных методах решения жестких задач при вычислении стадий применяется  $LU$ -разложение матрицы Якоби. В случае достаточно большой размерности исходной системы быстродействие алгоритма интегрирования фактически полностью определяется временем декомпозиции этой матрицы. Для повышения эффективности расчетов в ряде алгоритмов используется замораживание матрицы Якоби, то есть применение одной матрицы на нескольких шагах интегрирования [1]. Наиболее успешно этот подход применяется в многошаговых методах [2]. Не вызывает эта проблема особых трудностей и при построении алгоритмов интегрирования на основе других численных схем, если в них стадии вычисляются с участием матрицы Якоби в некотором итерационном процессе. В алгоритмах интегрирования на основе известных безытерационных численных схем, к которым относятся методы типа Розенброка [3] и их различные модификации [1], проблема замораживания более трудная. Следует отметить, что с точки зрения реализации безытерационные методы существенно проще алгоритмов на основе численных формул, в которых стадии вычисляются с применением итераций. Некоторым аналогом замораживания матрицы Якоби является применение в расчетах алгоритмов интегрирования на основе явных и  $L$ -устойчивых методов с автоматическим выбором численной схемы. В этом случае эффективность алгоритма может быть повышена за счет расчета переходных участков явным методом [4]. В качестве критерия выбора эффективной численной формулы естественно применять неравенство для контроля устойчивости [5].

Здесь на основе  $L$ -устойчивого (4,2)-метода и схемы Мерсона четвертого порядка точности построен алгоритм переменной структуры. Приведены результаты расчетов задачи проникновения помеченных радиоактивной меткой антител в пораженную опухолью ткань живого организма.

### 1. Исследование (4,2)-метода

Рассмотрим задачу Коши вида

$$y' = f(y), \quad y(t_0) = y_0, \quad t_0 \leq t \leq t_k, \quad (1)$$

где  $y$  и  $f$  – вещественные  $N$ -мерные вектор-функции,  $t$  – независимая переменная. Для решения задачи (1) рассмотрим численную формулу [6]

$$\begin{aligned} y_{n+1} &= y_n + p_1 k_1 + \dots + p_4 k_4, \quad D_n = E - ahf'_n, \\ D_n k_1 &= hf(y_n), \quad D_n k_2 = k_1, \end{aligned} \quad (2)$$

<sup>1</sup>Работа выполнена при финансовой поддержке РФФИ (проекты 11-01-00106 и 11-01-00224)

$$D_n k_3 = hf(y_n + \beta_{31}k_1 + \beta_{32}k_2) + \alpha_{32}k_2, \quad D_n k_4 = k_3 + \alpha_{42}k_2,$$

где  $h$  – шаг интегрирования,  $a, p_i, \beta_{ij}$  и  $\alpha_{ij}$  – числовые коэффициенты,  $k_i, 1 \leq i \leq 4$ , – стадии метода,  $E$  – единичная матрица,  $f'_n = \partial f / \partial y$  – матрица Якоби задачи (1). Подставим разложения  $k_i$  в виде рядов Тейлора в первую формулу (2). Полагая  $y_n = y(t_n)$  и сравнивая полученное представление с рядом Тейлора для точного решения, запишем условия четвертого порядка точности

$$\begin{aligned} 1) & p_1 + p_2 + (1 + \alpha_{32})p_3 + (1 + \alpha_{32} + \alpha_{42})p_4 = 1, \\ 2) & ap_1 + 2ap_2 + (a + \beta_{31} + \beta_{32} + 3a\alpha_{32})p_3 + \\ & + (2a + \beta_{31} + \beta_{32} + 4a\alpha_{32} + 3a\alpha_{42})p_4 = 0.5, \\ 3) & a^2 p_1 + 3a^2 p_2 + (a^2 + 2a\beta_{31} + 3a\beta_{32} + 6a^2\alpha_{32})p_3 + \\ & + (3a^2 + 3a\beta_{31} + 4a\beta_{32} + 10a^2\alpha_{32} + 6a^2\alpha_{42})p_4 = 1/6, \\ 4) & a^3 p_1 + 4a^3 p_2 + (a^3 + 3a^2\beta_{31} + 6a^2\beta_{32} + \\ & + 10a^3\alpha_{32})p_3 + (4a^3 + 6a^2\beta_{31} + 10a^2\beta_{32} + \\ & + 20a^3\alpha_{32} + 10a^3\alpha_{42})p_4 = 1/24, \\ 5) & (\beta_{31} + \beta_{32})^2(p_3 + p_4) = 1/3, \\ 6) & a(\beta_{31} + \beta_{32})(\beta_{31} + 2\beta_{32})(p_3 + p_4) = 1/8, \\ 7) & a(\beta_{31} + \beta_{32})^2(0.5p_3 + p_4) = 1/24, \\ 8) & (\beta_{31} + \beta_{32})^3(p_3 + p_4) = 1/4. \end{aligned} \quad (3)$$

Применяя метод (2) для решения тестовой задачи

$$y' = \lambda y, \quad y(0) = y_0, \quad t \geq 0,$$

имеем условие  $L$ -устойчивости численной формулы (2) вида

$$a(a - p_1) + (\beta_{31} - a)p_3 = 0.$$

Исследуя совместность этого соотношения и (3), запишем

$$\begin{aligned} p_1 &= \frac{76a^2 - 29a + 3}{27a^2}, \quad p_2 = \frac{146a^2 + 89a - 12}{27a^2}, \quad p_3 = \frac{32a - 4}{27a}, \\ p_4 &= \frac{4 - 16a}{27a}, \quad \beta_{31} = \frac{48a - 9}{32a}, \quad \beta_{32} = \frac{9 - 24a}{32a}, \\ \alpha_{32} &= \frac{-54a^2 + 57a - 12}{8a - 32a^2}, \quad \alpha_{42} = \frac{-864a^3 + 828a^2 - 288a + 36}{a(4 - 16a)^2}, \end{aligned} \quad (4)$$

где  $a$  определяется из необходимого условия  $L$ -устойчивости

$$24a^4 - 96a^3 + 72a^2 - 16a + 1 = 0.$$

Данное уравнение имеет четыре вещественных корня

$$\begin{aligned} a_1 &= 0.10643879214266, \quad a_2 = 0.22042841025921, \\ a_3 &= 0.57281606248213, \quad a_4 = 3.10031673511599. \end{aligned}$$

Для расчетов рекомендуется  $a = 0.57281606248213$ , потому что в этом случае схема (2) дополнительно является  $A$ -устойчивой.

Для контроля точности вычислений метода (2) четвертого порядка будем применять метод третьего порядка вида

$$y_{n+1,1} = y_n + b_1k_1 + b_2k_2 + b_3k_3 + b_4k_5,$$

где  $D_n k_5 = k_4$ , а  $k_i, 1 \leq i \leq 3$ , определены в (2). Нетрудно видеть, что требования третьего порядка имеют вид

$$\begin{aligned} b_1 + b_2 + (1 + \alpha_{32})b_3 + (1 + \alpha_{32} + \alpha_{42})b_4 &= 1, \\ ab_1 + 2ab_2 + (a + \beta_{31} + \beta_{32} + 3a\alpha_{32})b_3 + \\ (3a + \beta_{31} + \beta_{32} + 5a\alpha_{32} + 4a\alpha_{42})b_4 &= 1/2, \end{aligned}$$

$$\begin{aligned} & a^2 b_1 + 3a^2 b_2 + (a^2 + 2a\beta_{31} + 3a\beta_{32} + 6a^2\alpha_{32})b_3 + \\ & + (6a^2 + 4a\beta_{31} + 5a\beta_{32} + 15a^2\alpha_{32} + 10a^2\alpha_{42})b_4 = 1/6, \\ & (\beta_{31} + \beta_{32})^2 (b_3 + b_4) = 1/3, \end{aligned}$$

где коэффициенты  $a, \beta_{31}, \beta_{32}, \alpha_{32}$  и  $\alpha_{42}$  заданы в (4). Данная система линейна относительно параметров  $b_i, 1 \leq i \leq 4$ . При применении коэффициентов (4), имеем

$$\begin{aligned} b_1 &= 1.203100567018353, \quad b_2 = -6.552116304144386 \cdot 10^{-1}, \\ b_3 &= 7.115271884598151 \cdot 10^{-1}, \quad b_4 = -1.189345958672225 \cdot 10^{-1}. \end{aligned}$$

Теперь оценку ошибки  $\varepsilon_{n,0}$  можно вычислить по формуле  $\varepsilon_{n,0} = \|y_{n+1} - y_{n+1,1}\|$ , а при выборе шага проверить неравенство  $\varepsilon_{n,0} \leq \varepsilon$ , где  $\|\cdot\|$  – некоторая норма в  $R^N$ ,  $\varepsilon$  – точность расчетов.

Оценку максимального собственного числа  $\nu_{n,0} = h|\lambda_{n,\max}|$  матрицы Якоби системы (1), необходимую для перехода на явную формулу, оценим через ее норму по формуле

$$\nu_{n,0} = \|\partial f / \partial y\| = \max_{1 \leq i \leq N} \sum_{j=1}^N |\partial f_i(y_n) / \partial y_j|.$$

## 2. Контроль точности и устойчивости метода Мерсона

Одним из самых эффективных явных методов типа Рунге-Кутты четвертого порядка точности является метод Мерсона [7]

$$\begin{aligned} y_{n+1} &= y_n + \frac{1}{6}k_1 + \frac{2}{3}k_4 + \frac{1}{6}k_5, \\ k_1 &= hf(y_n), \quad k_2 = hf(y_n + \frac{1}{3}k_1), \quad k_3 = hf(y_n + \frac{1}{6}k_1 + \frac{1}{6}k_2), \\ k_4 &= hf(y_n + \frac{1}{8}k_1 + \frac{3}{8}k_3), \quad k_5 = hf(y_n + \frac{1}{2}k_1 - \frac{3}{2}k_3 + 2k_4). \end{aligned} \quad (5)$$

Пятое вычисление функции  $f$  не дает увеличение порядка до пятого, но позволяет расширить интервал устойчивости до 3.5 и оценить величину локальной ошибки  $\delta_{n,4}$  с помощью  $k_i$ , то есть

$$\delta_{n,4} = (2k_1 - 9k_3 + 8k_4 - k_5) / 30.$$

Для контроля точности будем применять неравенство  $\|\delta_{n,4}\| \leq 5\varepsilon^{5/4}$ , которое получено с учетом накопления глобальной ошибки из локальных ошибок [5]. Несмотря на то, что обоснование неравенства проведено на линейном уравнении, оно с достаточно высокой надежностью использовалось для решения нелинейных задач.

Теперь построим неравенство для контроля устойчивости. Применяя к разности  $k_3 - k_2$  формулу Тейлора с остаточным членом в форме Лагранжа первого порядка, имеем

$$k_3 - k_2 = h[\partial f(\mu_n) / \partial y](k_2 - k_1) / 6,$$

где вектор  $\mu_n$  вычислен в некоторой окрестности решения  $y(t_n)$ . Учитывая, что

$$k_2 - k_1 = h^2 f'_n f_n / 3 + O(h^3),$$

для контроля устойчивости (5) можно использовать неравенство

$$\nu_{n,4} = 6 \max_{1 \leq i \leq N} |(k_3 - k_2)_i / (k_2 - k_1)_i| \leq 3.5,$$

где числу 3.5 равна длина интервала устойчивости. Отметим, что по мнимой оси область устойчивости также ограничена числом 3.5 (рис. 1). Введем обозначения  $\varepsilon_{n,4} = \delta_{n,4} / 5$ . Тогда для контроля точности схемы (5) можно применять неравенство  $\varepsilon_{n,4} \leq \varepsilon^{5/4}$ , а для контроля устойчивости следующее –  $\nu_{n,4} \leq 3.5$ .

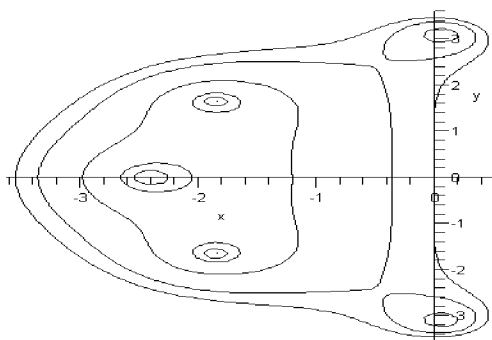


Рис. 1. Область устойчивости метода (5)

Так как оценка максимального собственного числа  $v_{n,4} = h|\lambda_{n,\max}|$  является грубой, то контроль устойчивости используется как ограничитель на размер шага интегрирования. В результате прогнозируемый шаг  $h_{n+1}$  вычисляется следующим образом. Новый шаг  $h^{ac}$  по точности определим по формуле  $h^{ac} = q_1 h_n$ , где  $h_n$  есть последний успешный шаг, а  $q_1$ , учитывая соотношение  $\varepsilon_{n,4} = O(h_n^5)$ , задается уравнением  $q_1^5 \varepsilon_{n,4} \leq \varepsilon$ . Шаг  $h^{st}$  по устойчивости зададим формулой  $h^{st} = q_2 h_n$ , где  $q_2$ , учитывая соотношение  $v_{n,4} = O(h_n)$ , определяется из равенства  $q_2 v_{n,4} = 3.5$ . Тогда прогнозируемый шаг  $h_{n+1}$  вычисляется по формуле

$$h_{n+1} = \max[h_n, \min(h^{ac}, h^{st})].$$

Данная формула стабилизирует шаг на участке установления решения, где определяющую роль играет устойчивость.

### 3. Алгоритм интегрирования

На основе построенных методов легко сформулировать алгоритм переменного шага, в котором на каждом шаге выбирается наиболее эффективная численная схема. Расчеты всегда начинаются явным методом потому, что в нем не используется матрица Якоби. Нарушение неравенства  $v_{n,4} \leq 3.5$  вызывает переход на  $L$ -устойчивую схему. Передача управления явным методам происходит в случае выполнения неравенства  $v_{n,0} \leq 3.5$ , где оценка  $v_{n,0}$  вычисляется через норму матрицы Якоби. Норма  $\|\xi\|$  в неравенствах для контроля точности вычисляется по формуле

$$\|\xi\| = \max_{1 \leq i \leq N} \{ |\xi_i| / (|y_n^i| + r) \}, \tag{6}$$

где  $i$  – номер компоненты,  $r$  – положительный параметр. Если по  $i$ -й компоненте решения выполняется неравенство  $\|y_n^i\| < r$ , то контролируется абсолютная ошибка  $re$ , в противном случае – относительная ошибка  $\varepsilon$ . Ниже построенный алгоритм переменного порядка и шага, а также с автоматическим выбором явной или  $L$ -устойчивой численной схемы будем называть РКМК4.

### 4. Результаты расчетов

Модель описывается системой двух дифференциальных уравнений в частных производных с начальными и граничными условиями. Это задача проникновения помеченных радиоактивной меткой антител в пораженную опухолью ткань живого организма [8]. Рассматривается система одномерных уравнений реакции-диффузии

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} - kuv, \quad \frac{\partial v}{\partial t} = -kuv, \tag{7}$$

которые возникают из химической реакции  $A+B \rightarrow C$  с константой скорости реакции  $k$ , где  $A$  – антитело с радиоактивной меткой, реагирующее с субстратом  $B$  – тканью, пораженной опухолью. Концентрации  $A$  и  $B$  обозначены через  $u$  и  $v$  соответственно. При выводе уравнений (7) предполагалось, что кинетика реакции описывается законом действующих масс, причем реагент  $A$  подвижен, тогда как реагент  $B$  неподвижен. Изучается полубесконечная пластина, внутри которой равномерно распределен субстрат  $B$ . Реагент  $A$ , попадая на поверхность пластины, начинает в нее проникать. Для моделирования проникновения уравнения (7) рассматриваются в полосе  $S_T = \{(x, t): 0 < x < \infty, 0 < t < T\}$  с начальными  $u(x, 0) = 0, v(x, 0) = v_0, x > 0$  и граничными  $u(0, t) = \Phi(t), 0 < t < T$  условиями, где  $v_0$  константа. Для численного решения переменная  $x$  преобразуется так, чтобы полубесконечная пластина преобразовалась в конечную. Такое преобразование обеспечивает специальное



семейство преобразований Мебиуса  $\zeta=x/(x+c)$ ,  $c>0$ . Это замена переменных преобразует  $S_T$  в прямоугольник  $\{(\zeta, t): 0<\zeta<1, 0<t<T\}$ . Через переменную  $\zeta$  задача (7) переписывается в виде

$$\frac{\partial u}{\partial t} = \frac{(\zeta-1)^4}{c^2} \frac{\partial^2 u}{\partial \zeta^2} + \frac{2(\zeta-1)^3}{c^2} \frac{\partial u}{\partial \zeta} - kuv, \quad \frac{\partial v}{\partial t} = -kuv \quad (8)$$

с начальными  $u(\zeta,0)=0$ ,  $v(\zeta,0)=v_0$ ,  $\zeta>0$  и граничными  $u(0, t)=\Phi(t)$ ,  $\partial u(1,t)/\partial \zeta=0$ ,  $0<t<T$  условиями. Последнее граничное условие получено из соотношения  $\partial u(\infty,t)/\partial x=0$ . Дискретизация производных по пространственным переменным с использованием метода прямых приводит к задаче Коши для системы обыкновенных дифференциальных уравнений. Для дискретизации применяется равномерная сетка  $\{\zeta_j\}$ ,  $\zeta_j=j \cdot \Delta \zeta$ ,  $\Delta \zeta=1/N$ ,  $1 \leq j \leq N$ . Через  $u_j$  и  $v_j$  обозначены аппроксимации  $u(\zeta_j, t)$  и  $v(\zeta_j, t)$ , соответственно. Очевидно, что  $u_j$  и  $v_j$  являются функциями от  $t$ . Дискретизации производных первого и второго порядков по пространственной переменной соответственно имеют вид

$$\frac{\partial u_j}{\partial \zeta} = \frac{u_{j+1} - u_{j-1}}{2\Delta \zeta}, \quad \frac{\partial^2 u_j}{\partial \zeta^2} = \frac{u_{j-1} - 2u_j + u_{j+1}}{(\Delta \zeta)^2}, \quad 1 \leq j \leq N.$$

Значения  $u_0$  и  $u_{N+1}$  получены из граничных условий, они имеют вид  $u_0=\Phi(t)$ ,  $u_{N+1}=u_N$ . Полагая  $y=(u_1, v_1, u_2, v_2, \dots, u_N, v_N)^T$  и  $T=20$ , эта задача имеет вид

$$y' = f(t, y), \quad y(0) = g, \quad y \in R^{2N}, \quad 0 \leq t \leq 20, \quad (9)$$

где  $N$  – параметр, а функция  $f$  определяется формулами

$$f_{2j-1} = \alpha_j \frac{y_{2j+1} - y_{2j-3}}{2\Delta \zeta} + \beta_j \frac{y_{2j-3} - 2y_{2j-1} + y_{2j+1}}{(\Delta \zeta)^2} - ky_{2j-1}y_{2j}, \quad f_{2j} = -ky_{2j}y_{2j-1},$$

где  $g=(0, v_0, 0, v_0, \dots, 0, v_0)^T$ ,  $\alpha_j=2(j \cdot \Delta \zeta - 1)^3 c^2$ ,  $\beta_j=(j \cdot \Delta \zeta - 1)^4 c^2$ ,  $\Delta \zeta=1/N$ ,  $1 \leq j \leq N$ ,  $y_{-1}(t)=\Phi(t)$ ,  $y_{2N+1}=y_{2N-1}$ . Функция  $\Phi(t)=2$  при  $0<t \leq 5$  и  $\Phi(t)=0$  при  $5<t \leq 20$ , то есть  $\Phi(t)$  имеет разрыв первого рода в точке  $t=5$ . Согласно [8] подходящими значениями для параметров  $k$ ,  $v_0$  и  $c$  являются  $k=100$ ,  $v_0=1$  и  $c=4$ . Расчеты проводились при  $N=400$ , то есть система (9) состоит из 800 уравнений. Задача о нахождении разрыва функции  $\Phi(t)$  при  $t=5$  возлагалась на алгоритм управления шагом. Решение задачи (9) приведено на рис. 2.

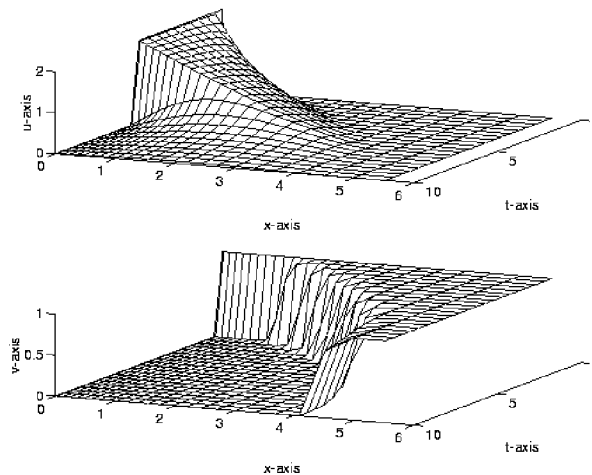


Рис. 2. Решение задачи (7)

Расчеты проводились с двойной точностью. Параметр  $r$  в формуле (6) выбирался таким образом, чтобы фактическая точность вычисления решения совпадала с задаваемой точностью. Сравнение проводилось с приближенным решением, которое получено при задаваемой точности вычислений  $\varepsilon=10^{-9}$  различными методами. Матрица Якоби вычислялась численно. В табл. 1 приведены результаты расчетов при различных значениях задаваемой точности вычислений  $\varepsilon$ . Вычислительные затраты приведены в форме  $if(ij)$ , где через  $if$  и  $ij$  обозначены соответственно число вычислений правой части и декомпозиций матрицы Якоби на интервале интегрирования. В табл. 1 используются обозначения: RK4 – метод Мерсона, RK4st – метод Мерсона с контролем устойчивости, МК4 –  $L$ -устойчивый метод, РКМК4 – алгоритм с автоматическим выбором численной схемы.

Таблица 1

## Результаты расчетов

Метод / $\varepsilon$	RK4	RK4st	MK4	RKMK4
$10^{-2}$	1 017 921	883 561	39 560(49)	35 286(34)
$10^{-3}$	1 018 024	886 914	62 918(78)	44 923(46)
$10^{-4}$	1 018 177	889 604	76 717(95)	66 845(74)
$10^{-5}$	1 018 701	882 836	92 165(114)	83 984(86)
$10^{-6}$	1 019 268	916 280	103 687(128)	96 205(99)

Из сравнения результатов расчетов RK4 и RK4st следует, что на данной задаче контроль устойчивости приводит к сокращению вычислительных затрат более чем на 10%. Это является следствием устранения некоторых повторных вычислений решения, возникающих за счет неустойчивости численной схемы. На других задачах преимущество RK4st перед RK4 может достигать 50% [5]. Из сравнения результатов расчетов MK4 и RKMK4 следует сокращение числа декомпозиций матрицы Якоби, что является следствием расчета некоторых переходных участков по явной численной формуле.

## Заключение

В RKMK4 с помощью признака можно задавать различные режимы расчета: 1) явным методом с контролем или без контроля устойчивости; 2)  $L$ -устойчивым методом с аналитической или численной матрицей Якоби; 3) с автоматическим выбором численной схемы. Все это позволяет применять данный алгоритм для решения как жестких, так и нежестких задач. При расчетах с автоматическим выбором численной схемы вопрос о том, является ли задача жесткой или нет, перекладывается на алгоритм интегрирования.

## Литература

1. Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. – М.: Мир, 1999. – 685 с.
2. Byrne G.D., Hindmarsh A.C. ODE solvers: a review of current and coming attractions // J. of Comput. Physics. 1987. №70. P. 1–62.
3. Rosenbrock H.H. Some general implicit processes for the numerical solution of differential equations // Computer. 1963. №5. P. 329–330.
4. Новиков А.Е., Новиков Е.А. Численное решение жестких задач с небольшой точностью // Математическое моделирование. 2010. Т.22, №1. С. 46–56.
5. Новиков Е.А. Явные методы для жестких систем. – Новосибирск: Наука, 1997. – 197 с.
6. Новиков Е.А., Шитов Ю.А., Шокин Ю.И. Одношаговые безытерационные методы решения жестких систем // ДАН СССР. 1988. Т. 301, №6. С. 1310–1314.
7. Merson R.H. An operational methods for integration processes // Proc. of Symp. on Data Processing. Weapons Research Establishment, Salisbury, Australia. 1957. P. 331.
8. Mazzia F., Iavernaro F. Test Set for Initial Value Problem Solvers // Department of Mathematics, University of Bari, August 2003. URL: <http://www.dm.uniba.it/~testset>.

Новиков Антон Евгеньевич, ассистент кафедры Сибирского федерального университета, тел. (391)2495979, e-mail: [aenovikov@bk.ru](mailto:aenovikov@bk.ru)

Шайдуров Владимир Викторович, член-корреспондент РАН, директор Института вычислительного моделирования СО РАН, тел. (391)290-74-76, e-mail: [shaidurov04@mail.ru](mailto:shaidurov04@mail.ru)

Novikov Anton Evgen'evich, teaching assistant of the Siberian Federal University, tel. +7 (391)2495979, e-mail: [aenovikov@bk.ru](mailto:aenovikov@bk.ru)

Shaidurov Vladimir Viktorovich, corresponding member of RAS, director of Institute of Calculation Mathematics of SB RAS, tel. (391)290-74-76, e-mail: [shaidurov04@mail.ru](mailto:shaidurov04@mail.ru)

© В.А. Русанов, Л.В. Антонова, А.В. Данеев, А.С. Миронов

## О ФОРМАЛЬНО РАЗРЕШИМЫХ РЕАЛИЗАЦИЯХ ДИНАМИЧЕСКИХ СИСТЕМ С МИНИМАЛЬНОЙ ОПЕРАТОРНОЙ НОРМОЙ<sup>1</sup>

Проведено изучение необходимых и достаточных условий существования нелинейных дифференциальных реализаций управляемых бихевиористических систем (динамических систем Я. Виллемса) в классе квазилинейных дифференциальных уравнений с минимальной операторной нормой в равномерно выпуклом банаховом пространстве.

**Ключевые слова:** динамическая система, дифференциальная реализация с минимальной операторной нормой.

© V.A. Rusanov, L.V. Antonova, A.V. Daneev, A.S. Mironov

## ON THE FORMAL REALIZATIONS OF SOLVABLE DYNAMICAL SYSTEMS WITH MINIMAL OPERATOR NORM<sup>1</sup>

The study of necessary and sufficient conditions for the existence of nonlinear differential implementations of managed behavioral sciences systems (dynamical systems, J. Willems) in the class of quasi-linear differential equations with minimal operator norm in uniformly convex Banach space.

**Key words:** dynamical systems, differential implementation with minimal operator norm.

### Введение

Теория реализации изучает [1, с. 21] проблему апостериорного математического моделирования уравнений состояния прогностической динамической системы в её наиболее чистом виде, когда модель рассматривается лишь с точки зрения соответствия (или несоответствия) некоторому фиксированному набору экспериментальных данных [2], представленных пучком вектор-функций «траектория, программное управление, нелинейное позиционное управление» – бихевиористическая модель [3] динамической системы (*D*-системы); при этом подходе все остальные возможные свойства исследуемой модели остаются в стороне. В данном контексте в [4,5] были изучены характеристические критерии формальной разрешимости задачи дифференциальной реализации на произвольных пучках (фиксированных семействах) управляемых динамических процессов в сепарабельных банаховых пространствах, в частности, равномерно выпуклых [5]. В них рассматривались понятия и отношения наиболее общего вида. Дальнейшая задача заключается в более глубоком конструктивном изучении дифференциальной реализации; желательно привести общие свойства разрешимости в связь с моделями более конкретными, допускающими «непосредственное» изучение, при этом либо понять, что к означенной задаче применим уже существующий аппарат, либо изобрести новый.

Обсуждая роль математики в формулировании физических законов и выводе из них прогностических следствий, необходимо отметить часто возникающее соответствие между глубиной физической теории (когда лежащие в её основе физические идеи и представления чрезвычайно тонки и глубоки) и степенью элегантности ее математического аппарата. И хотя бесполезно было бы пытаться сколько-нибудь полно описать взаимодействие между математикой и науками из ряда прогностического моделирования, остановимся, однако, на одном важном аспекте этого взаимодействия, представляющем значительный методологический интерес.

Внешний мир настолько сложен, что ученый-естествоиспытатель бывает доволен, если ему удастся уловить и понять хотя бы некоторые самые простые из присущих миру закономерностей. Для этого он вводит упрощенные и идеализированные модели (в частности, и квазилинейные), освобожденные от маловажных и усложняющих дело подробностей-деталей и отражающие, как он надеется, наиболее существенные свойства рассматриваемых физических объектов. В ходе этого процесса, грубо говоря, дело обстоит так: вопрос о выборе класса модели решает ученый-естествоиспытатель, после этого свою роль выполняет МАТЕМАТИКА, позволяющая дедуктивно

---

<sup>1</sup> Работа выполнена при поддержке Программы фундаментальных исследований № 15 Отделения энергетики, машиностроения, механики и процессов управления РАН (проект № 2.5).

выводить заключения уже только на основе предложенного модельного класса. Все это достаточно хорошо известно и вряд ли требовало дальнейшего обсуждения, если бы не обстоятельство «множественности» моделей, которое может возникнуть в обстоятельствах, когда апостериорных данных не достаточно, чтобы определить (в означенном классе) модель *однозначно*. Исходя из этих соображений, кажется оправданным классифицировать эти модели на основе решения вспомогательной задачи *оптимального* моделирования на базе формального критерия, интерпретированного некоторым ясным физическим контекстом.

**1. Постановка задачи дифференциальной реализации нелинейной бихевиористической D-системы с минимальной операторной нормой**

Везде далее  $(X, \|\cdot\|_X)$ ,  $(Y, \|\cdot\|_Y)$  и  $(Z, \|\cdot\|_Z)$  – вещественные сепарабельные банаховы пространства,  $L(Y, X)$  – банахово пространство с операторной нормой  $\|\cdot\|_{L(Y, X)}$  всех линейных непрерывных операторов, действующих из  $Y$  в  $X$  (аналогично  $L(X, X)$ ,  $\|\cdot\|_{L(X, X)}$  и  $L(Z, X)$ ,  $\|\cdot\|_{L(Z, X)}$ ),  $T := [t_0, t_1]$  – отрезок числовой прямой  $R$  с мерой Лебега  $\mu$  и  $\nu$  – положительная мера, абсолютно непрерывная относительно  $\mu$  и определенная на  $\sigma$ -алгебре  $\mathcal{A}$  всех  $\nu$ -измеримых (лебеговски пополненных [6, с. 73]) подмножеств из  $T$ , через  $p, q \in (1, \infty)$  обозначим сопряженные числа  $p^{-1} + q^{-1} = 1$  [6, с. 145].

Пусть  $(B, \|\cdot\|_B)$  – банахово пространство,  $\mathcal{L}_r(T, \nu, B)$ ,  $r \in (1, \infty)$  – пространство всех интегрируемых (по Бохнеру [7, с. 189]) отображений  $f: T \rightarrow B$  с нормой  $\|f\|_{B, L^r} := (\int_T \|f(\tau)\|_B^r \nu(d\tau))^{1/r}$ . Как обычно,  $L_r(T, \nu, B)$  – банахово фактор-пространство классов  $\nu$ -эквивалентности в  $\mathcal{L}_r(T, \nu, B)$ , через  $AC(T, B) \mathfrak{L}_1(T, \mu, B)$  – линейное множество всех абсолютно непрерывных функций (относительно меры  $\mu$ ).

Определим  $(B, \|\cdot\|_B)$ -пространства с дополнительной структурой:

Определение 1 [7, с. 182]. *Банахово пространство  $(B, \|\cdot\|_B)$  называют равномерно выпуклым, если каково бы ни было  $\varepsilon > 0$ , существует такое  $\delta = \delta(\varepsilon) > 0$ , что из  $\|x\|_B \leq 1$ ,  $\|y\|_B \leq 1$  и  $\|x - y\|_B \geq \varepsilon$  ( $x, y \in B$ ) следует неравенство  $\|x + y\|_B \leq 2(1 - \delta)$ .*

*Замечание 1. а) Есть перефразировка: банахово  $(B, \|\cdot\|_B)$  равномерно выпукло, если из  $\|x\|_B \leq 1 + \varepsilon$ ,  $\|y\|_B \leq 1 + \varepsilon$ ,  $\|(x + y)/2\|_B > 1$  следует  $\|x - y\|_B \rightarrow 0$  вместе с  $\varepsilon \rightarrow 0$ .*

*б) Любое гильбертово пространство  $(G, \|\cdot\|_G)$  будет a fortiori равномерно выпукло согласно [7, с. 182] характеристической формулы «предгильбертовости»*

$$\|x + y\|_G^2 + \|x - y\|_G^2 = 2(\|x\|_G^2 + \|y\|_G^2), \forall x, y \in G.$$

*в) Всякое равномерно выпуклое пространство рефлексивно [7, с. 182].*

Выделим к рассмотрению дифференциальные модели класса

$$dx(t)/dt = A(t)x(t) + B(t)u(t) + B^\#(t)u^\#(x(t)), \tag{1}$$

где  $x \in AC(T, X)$  – решение Каратеодори ( $K$ -решение),  $u \in L_q(T, \mu, Y)$  и  $u^\#(x) \in L_q(T, \mu, Z)$  – программное и нелинейное позиционное управления,  $(A, B, B^\#) \in L_p(T, \mu, L(X, X)) \times L_p(T, \mu, L(Y, X)) \times L_p(T, \mu, L(Z, X))$ ; в целях удобства вектор-функцию  $(x, u, u^\#(x))$  из (1) тоже будем называть  $K$ -решением, а тройку оператор-функций  $(A, B, B^\#)$ , согласно терминологии из [5], –  $(A, B, B^\#)_p$ -моделью дифференциальной системы (1).

Теперь введем несколько важных конструкций. Обозначим через

$$L_p := L_p(T, \mu, L(X, X)) \times L_p(T, \mu, L(Y, X)) \times L_p(T, \mu, L(Z, X))$$

банахово пространство классов  $\mu$ -эквивалентности всех  $(A, B, B^\#)_p$ -моделей (упорядоченных троек оператор-функций из уравнений (1)) с нормой

$$\begin{aligned} \|(A, B, B^\#)\|_L &:= \\ &= (\int_T (\|A(\tau)\|_{L(X, X)}^p + \|B(\tau)\|_{L(Y, X)}^p + \|B^\#(\tau)\|_{L(Z, X)}^p) \mu(d\tau))^{1/p}. \end{aligned}$$

Через  $(H_q, \|\cdot\|_H)$  обозначим пространство-произведение (с нормой)

$$H_q := L_q(T, \mu, X) \times L_q(T, \mu, Y) \times L_q(T, \mu, Z),$$

$$\|(g, w, v)\|_H := \left( \int_T (\|g(\tau)\|_X^q + \|w(\tau)\|_Y^q + \|v(\tau)\|_Z^q) \mu(d\tau) \right)^{1/q}, \quad (g, w, v) \in H_q,$$

которое, как полное (в силу конструкции нормы  $\|\cdot\|_H$ ), является банаховым. Далее, через  $(L(H_q, X), \|\cdot\|_{L(H_q, X)})$  обозначим банахово пространство с операторной нормой всех линейных непрерывных операторов, действующих из  $H_q$  в  $X$ .

Пусть  $(A, B, B^\#) \in L_p$ . Рассмотрим оператор  $\xi: H_q \rightarrow X$ , имеющий представление

$$\xi(g, w, v) := \int_T (A(\tau)g(\tau) + B(\tau)w(\tau) + B^\#(\tau)v(\tau)) \mu(d\tau), \quad (g, w, v) \in H_q; \quad (2)$$

ясно, что  $\xi \in L(H_q, X)$ ; по терминологии [5] оператор  $\xi - \xi_p$ -модель. Для обратного утверждения: « $\zeta \in L(H_q, X) \Rightarrow$  оператор  $\zeta$  имеет аналитическое представление (2)», требуются *mutatis mutandis* уточнения и рассуждения (см. лемму 2 [5]).

В задачах апостериорного моделирования уравнений динамики сложных управляемых динамических процессов понимание физической природы функционирования исследуемой  $D$ -системы, представленной процессами «вход-выход», не всегда достаточно для определения структуры её математической модели; например, допустима ли линеаризация уравнений модели на основе анализа полученных экспериментальных данных. В такой постановке возникает следующий методологический вопрос: какой из возможных математических структур модели следует отдать предпочтение при имеющихся входных и выходных данных? На самом деле, в большинстве практических случаев, ответ на означенный вопрос должен быть одним из первых шагов в построении математической модели  $D$ -системы. Следуя этой парадигме, началом в исследовании дифференциальной реализации системы часто служит следующая обратная задача системного анализа [4,5]:

*Разрешимость дифференциальной реализации на пучке процессов<sup>2</sup>*: пусть

$$u^\#(\cdot): AC(T, X) \rightarrow L_q(T, \mu, Z), \quad q \in \{1, \infty\},$$

$$N \in \Pi_{u^\#} := \{(x, u, v) \in AC(T, X) \times L_q(T, \mu, Y) \times L_q(T, \mu, Z) : (x, u, v) = (x, u, u^\#(x))\} \quad (3)$$

где  $N$  – фиксированное экзогенное поведение типа «вход-выход» исследуемой  $D$ -системы с нелинейным позиционным управлением  $u^\#(x)$ , заданным *a priori*. Определить необходимые и достаточные условия, при которых пучок динамических процессов  $N$  представляет  $K$ -решения некоторого дифференциального уравнения (1); ограничений на  $\text{Card } N$  (мощность пучка  $N$ ) не накладываем (например, допустимо положение, когда  $\text{Card } N \geq \aleph_0$  – алеф нуль).

При построении дифференциальной реализации для фиксированного семейства процессов (3), как правило (даже для конечномерных систем (2)), имеем дело не с одной системой уравнений реализации (2), а с целым семейством систем [10], что выдвигает задачи построения «оптимальных реализаций» по некоторым формальным критериям (речь не идет о *минимальных* реализациях [11, с. 165]). Данные постановки допускают несколько математических трактовок, а именно, – можно исходить из задачи «оптимальной реализации» в структуре пространства  $L(H_q, X)$  по критерию нормы  $\|\cdot\|_{L(H_q, X)}$  (см. раздел 3), или пространства  $L_p$  по критерию нормы  $\|\cdot\|_L$  (раздел 4); см. так же замечание 3<sup>о</sup> и пример 1 [12]. Начнем с задачи существования реализации с *минимальной операторной нормой* в пространстве  $(L(H_q, X), \|\cdot\|_{L(H_q, X)})$ .

Перейдем к деталям. Банахово пространство  $X$  по геометрии локально выпукло, следовательно (т.к. пространство сопряженное  $X$  разделяет на  $X$  точки), интегральный оператор  $\Gamma: L_p \rightarrow L(H_q, X)$ , осуществляющий согласно (2) соответствие

$$\Gamma(A, B, B^\#) := \xi, \quad (4)$$

<sup>2</sup> Означенная постановка дифференциальной реализации  $D$ -системы не исключают методологического положения, когда закон  $u^\#(x)$  детерминируется не по принципу «state feedback», а характеризует существенную «нелинейную компоненту» в уравнениях динамики (1), моделируемых *a posteriori* [8], опираясь (и развивая) на результаты теории геометрии  $u^\#(x)$ -поверхностей [9].

линейный гомеоморфизм *ipso facto* (лемма 2 [5]) между линейными множествами всех  $(A, B, B^\#)_p$ - и  $\xi_p$ - моделями (дифференциальными и интегральными моделями), при этом  $(A, B, B^\#)_p$ - и  $\xi_p$ - модели (4) будем называть  *$\Gamma$ -изоморфными* [13, с. 124]. Здесь уместно отметить, что [13, с. 124] эквивалентные нормы  $\|\cdot\|_X, \|\cdot\|^*$  на  $X$   $\alpha_1\|\cdot\|_X \leq \|\cdot\|^* \leq \alpha_2\|\cdot\|_X, \alpha_1, \alpha_2 > 0 \in R$  не определены однозначно, но тоже «производят» изоморфизм  $(A, B, B^\#)_p$ - и  $\xi_p$ - моделей для  $(L_p, \|\cdot\|_L)$  и  $(L(H_q, X), \|\cdot\|_{L(H, X)})$ ; как правило, обычно в каждом случае для рассматриваемого изоморфизма существует естественный (прагматический) выбор нормы  $\|\cdot\|^*$ , и он может быть «по-своему хорош» (по крайней мере, практически оправдан в некоторой физической интерпретации).

**Определение 2.** Для динамических процессов (3) дифференциальную реализацию (1) с  $\xi_p$ - моделью  $\xi_{\min} \in L(H_q, X)$ , для которой при  $(X, \|\cdot\|^*)$  имеет место

$$\|\xi_{\min}\|_{L(H, X)} = \min \{ \|\xi\|_{L(H, X)}; \xi - \text{оператор (2)} \\ \text{некоторой реализации (1) семейства процессов (3)} \},$$

назовем реализацией оптимальной по критерию операторной нормы  $\|\cdot\|_{L(H, X)}$ , при этом  $(A, B, B^\#)_p$ - модель  $\Gamma^{-1}(\xi_{\min})$  назовем  $\|\cdot\|_{L(H, X)}$ -оптимальной с  $(X, \|\cdot\|^*)$ .

Стоит подчеркнуть, что можно провести четкую физическую границу между оптимизацией по критерию нормы  $\|\cdot\|_{L(H, X)}$  или по критерию нормы  $\|\cdot\|_L$ ; несомненно, при желании здесь тоже можно найти новые темы для размышления.

## 2. Распространение линейных ограниченных операторов с сохранением линейности и нормы

Все выдающиеся достижения математического моделирования связаны с тем, что оно позволяет абстрагировать определенные свойства наблюдаемых объектов и наблюдаемых отношений между ними и чисто логическим путем выводить новые свойства и новые отношения, которые можно затем проверить экспериментом. При этом в самой математике новая теория обычно начинается с того, что постулируется ряд новых математических свойств. Как скоро убедимся, вопрос о существовании  $\xi_{\min}$ -модели для  $N$  решается в зависимости не только от аналитической структуры пучка  $N$ , но и от метрических структур пространств «входных» –  $H_q$  и «выходных» –  $X$  сигналов пространства операторных  $\xi$ -моделей  $L(H_q, X)$ .

**Определение 3** [13, с. 246]<sup>3</sup>. Система множеств называется сцепленной, если каждая пара множеств из этой системы имеет непустое пересечение, при этом будем говорить, что линейное пространство  $x$  есть пространство структуры  $\mathfrak{R}$ , если  $(x, \|\cdot\|_{\mathfrak{R}})$  является нормированным пространством, каждое сцепленное семейство замкнутых шаров которого имеет непустое пересечение.

Заметим [13, с. 246], что всякое нормированное пространство  $x$  типа  $\mathfrak{R}$  – банахово. Действительно пусть  $\{x_n\}$  – сходящаяся в себе последовательность элементов пространства  $x$ . Обозначим через  $r_n$  значение  $\sup \{ \|x_m - x_n\|_{\mathfrak{R}}; m \geq n \}$  и рассмотрим семейство  $\{Sr_n(x_n)\}_{n=1, 2, \dots}$  замкнутых шаров радиусов  $r_n$  с центрами в  $x_n$ . Это сцепленное семейство, поскольку при  $m > n$  будет  $\|x_m - x_n\|_{\mathfrak{R}} \leq r_n$  и, следовательно,  $x_n \in Sr_n(x_n) \cap Sr_m(x_m)$ . Если  $x$  – пространство типа  $\mathfrak{R}$ , то существует элемент  $x \in x$ , входящий в каждый из шаров  $Sr_n(x_n)$ , т.е. имеем  $\|x - x_n\|_{\mathfrak{R}} \leq r_n (n=1, 2, \dots)$  и поскольку  $r_n \rightarrow 0$ , то  $x_n \rightarrow x$ . Тем самым полнота пространства  $x$  установлена.

Следующие два предложения установлены Нахбиным [14].

<sup>3</sup> Простейшим примером пространства типа  $\mathfrak{R}$  является вещественная числовая прямая. Действительно, замкнутый шар в этом пространстве представляет собой замкнутый промежуток. Рассмотрим поэтому сцепленное множество  $\{[a_\alpha, b_\alpha]\}$  ( $\alpha \in A$ ) замкнутых промежутков и проверим, что это множество имеет непустое пересечение. Фиксируем для этого некоторое  $\alpha^* \in A$ . Поскольку промежутки  $[a_\alpha, b_\alpha]$  и  $[a_{\alpha^*}, b_{\alpha^*}]$  имеют общие точки, каково бы ни было  $\alpha \in A$ , то  $a_{\alpha^*} \leq b_\alpha$  ( $\alpha \in A$ ). Следовательно,  $a_{\alpha^*} \leq \inf \{ b_\alpha; \alpha \in A \}$ . А так как  $\alpha^* \in A$  произвольно, то отсюда  $a^* = \sup \{ a_\alpha; \alpha \in A \} \leq \inf \{ b_\alpha; \alpha \in A \} = b^*$ . Таким образом, любая точка промежутка  $[a^*, b^*]$  входит в каждый из промежутков  $[a_\alpha, b_\alpha]$ , что в конечном итоге и доказывает исходное утверждение.

Предложение 1.<sup>4</sup> Пусть  $(H, \|\cdot\|_H)$  – нормированное пространство и  $\Omega^\#$  – линейное множество, содержащееся в  $H$ . Пусть, далее,  $\xi^\#$  – линейный непрерывный оператор, отображающий  $\Omega^\#$  в нормированное пространство  $(X, \|\cdot\|_X)$  типа  $\mathfrak{R}$ . Тогда существует линейное непрерывное распространение  $\xi^+$  оператора  $\xi^\#$ , отображающее пространство  $H$  в пространство  $X$ , причём выполняется

$$\|\xi^+\|_{L(H, X)} = \|\xi^\#\|_{L(\Omega^\#, X)}.$$

**Доказательство.** Линейное множество  $\Omega^*$ , состоящее из всех элементов вида  $\alpha\omega^* + \omega$ , где  $\omega^*$  – фиксированный элемент из  $H$ ,  $\alpha \in \mathfrak{R}$ ,  $\omega \in \Omega^\#$  назовём элементарным расширением линейного множества  $\Omega^\#$ . Покажем вначале, что условия предложения позволяют всякий оператор  $\xi^\# \in \mathcal{L}(\Omega^\#, X)$  линейно распространить с сохранением нормы на любое элементарное расширение  $\Omega^*$  множества  $\Omega^\#$ ; разумеется, что речь идёт о собственном расширении, – когда  $\omega^* \notin \Omega^\#$ .

Ясно, что если распространение  $\xi^* \in \mathcal{L}(\Omega^*, X)$  оператора  $\xi^\# \in \mathcal{L}(\Omega^\#, X)$  на элементарное расширение  $\Omega^*$  существует, то оно определяется указанием в пространстве  $X$  элемента  $x^* = \xi^*(\omega^*)$ , при этом, очевидно, для вектора  $x^*$  должна выполняться следующая цепочка транзитивных отношений:

$$\begin{aligned} \|x^* - \xi^\#(\omega)\|_X &= \|\xi^*(\omega^*) - \xi^*(\omega)\|_X \leq \\ &\leq \|\xi^*\|_{L(\Omega^*, X)} \|\omega^* - \omega\|_H = \|\xi^\#\|_{L(\Omega^\#, X)} \|\omega^* - \omega\|_H. \end{aligned}$$

Таким образом, необходимым условием существования оператора  $\xi^* \in \mathcal{L}(\Omega^*, X)$  является наличие общей точки  $x^*$  у всех замкнутых  $r_\omega$ -шаров  $Sr_\omega(x_\omega)$  с центрами, соответственно, в точках  $x_\omega = \xi^\#(\omega)$  и радиусами  $r_\omega = \|\xi^\#\|_{L(\Omega^\#, X)} \|\omega^* - \omega\|_H$ , где  $\omega \in \Omega^\#$ . Теперь покажем, что это условие является также и достаточным.

Пусть  $U = \{Sr_\omega(x_\omega)\}_{\omega \in \Omega^\#}$  и предположим, что имеется точка  $x^*$ , принадлежащая всем  $r_\omega$ -шарам семейства  $U$ . Условимся, что  $\xi^*(\omega^*) = x^*$  и  $\xi^*(\omega') = \alpha x^* + \xi^\#(\omega)$  для  $\omega' = \alpha\omega^* + \omega \in \Omega^*$ . Далее, так как при  $\alpha \neq 0$  справедлива цепочка отношений

$$\|\xi^*(\omega')\|_X = |\alpha| \|x^* - \xi^\#(\omega/\alpha)\|_X \leq |\alpha| r_{\omega/\alpha} = |\alpha| \|\xi^\#\|_{L(\Omega^\#, X)} \|\omega^* - \omega/\alpha\|_H = \|\xi^\#\|_{L(\Omega^\#, X)} \|\omega'\|_H,$$

то  $\xi^*$  – линейный непрерывный оператор, переводящий  $\Omega^*$  в  $\{\alpha x^* + \xi^\#(\omega) : \alpha \in \mathfrak{R}, \omega \in \Omega^\#\}$ , причём  $\|\xi^*\|_{L(\Omega^*, X)} \leq \|\xi^\#\|_{L(\Omega^\#, X)}$ . Обратное неравенство очевидно. Поэтому окончательно  $\|\xi^*\|_{L(\Omega^*, X)} = \|\xi^\#\|_{L(\Omega^\#, X)}$ . Теперь остаётся показать, что шары семейства  $U$  имеют непустое пересечение, для чего (принимая во внимание посылку предложения, касающуюся задания типа  $\mathfrak{R}$  пространства  $X$ ) достаточно проверить, что  $U$  – сцепленное семейство. Последнее очевидно, рассмотрим два произвольных шара  $Sr_{\omega'}(x_{\omega'})$  и  $Sr_{\omega''}(x_{\omega''})$  из  $U$ . Ясно, что для этих шаров выполняется

$$\begin{aligned} r_{\omega'} + r_{\omega''} &= \|\xi^\#\|_{L(\Omega^\#, X)} (\|\omega^* - \omega'\|_H + \|\omega^* - \omega''\|_H) \geq \\ &\geq \|\xi^\#\|_{L(\Omega^\#, X)} \|(\omega^* - \omega') - (\omega^* - \omega'')\|_H \geq \|\xi^\#\|_{L(\Omega^\#, X)} \|\omega' - \omega''\|_H = \|x_{\omega'} - x_{\omega''}\|_X, \end{aligned}$$

т.е. сумма их радиусов не меньше расстояния между их центрами, следовательно, шары  $Sr_{\omega'}(x_{\omega'})$  и  $Sr_{\omega''}(x_{\omega''})$  пересекаются<sup>5</sup>. Последнее (в виду произвольности сделанного выбора шаров из  $U$ ) означает, что семейство  $U$  сцепленное.

Вторую часть доказательства можно провести одним из канонических способов трансфинитной индукции<sup>6</sup> (выбранным по вкусу). Можно воспользоваться либо постулатом Цермело, либо лем-

<sup>4</sup> Необходимость введения в рассмотрение пространств структуры  $\mathfrak{R}$  обусловлена тем, что аналитический аппарат  $\xi_{\min}$ -моделей можно строить в рамках теории о распространении линейного оператора с сохранением его нормы [14], как плодотворного аналога классической теоремы Хана–Банаха. Ниже сформулируем и докажем предложения 2, 3 и тем самым в некотором смысле *полностью* охарактеризуем линейное расширение оператора  $\xi_{\min}$ -моделей с инвариантной нормой.

<sup>5</sup> Если  $S_r(x')$  и  $S_{r'}(x'')$  – два замкнутых шара пространства  $X$  и  $r + r' \geq \|x' - x''\|_X$ , то их пересечение непусто; например, имеет место включение  $r'x' / (r' + r'') + r x'' / (r + r') \in S_r(x') \cap S_{r'}(x'')$ .

мой Цорна, либо принципом максимальности Хаусдорфа [15, с. 54]; ниже предпочтён последний способ.

Пусть  $\Sigma$  – семейство всех упорядоченных пар  $(\Omega, \xi)$ , где  $\Omega$  – линейное множество в пространстве  $H$ , содержащее  $\Omega^\#$ , а  $\xi \in \mathcal{L}(\Omega, x)$  – такой оператор, что  $\xi = \xi^\#$  на  $\Omega^\#$  и  $\|\xi\|_{L(\Omega, x)} = \|\xi^\#\|_{L(\Omega^\#, x)}$ . Введём в  $\Sigma$  частичное упорядочение, считая, что  $(\Omega', \xi') \leq (\Omega'', \xi'')$  (для  $(\Omega', \xi'), (\Omega'', \xi'') \in \Sigma$ ), если  $\Omega' \subset \Omega''$  и  $\xi' = \xi''$  на  $\Omega'$ . В соответствии с принципом Хаусдорфа в  $\Sigma$  существует максимальное гнездо  $\Sigma^+$ .

Пусть  $D$  – множество всех таких линейных множеств  $\Omega$  пространства  $H$ , что  $(\Omega, \xi) \in \Sigma^+$  для некоторого  $\xi \in \mathcal{L}(\Omega, x)$ . Тогда  $D$  линейно упорядоченно относительно теоретико-множественного включения и, таким образом,  $\Omega^+ = \cup \{ \Omega : \Omega \in D \}$  – линейное множество. Если  $\omega \in \Omega^+$ , то, очевидно,  $\omega \in \Omega$ , для некоторого  $\Omega \in D$ . Далее, положим  $\xi^+(\omega) = \xi(\omega)$ , где  $\xi \in \mathcal{L}(\Omega, x)$  – оператор, составляющий вместе с  $\Omega$  пару  $(\Omega, \xi) \in \Sigma^+$ . Несложно проверить, что оператор  $\xi^+$  корректно определён на  $\Omega^+$ , к тому же он линейен, совпадает с  $\xi^\#$  на  $\Omega^\#$  и удовлетворяет на  $\Omega^+$  соотношению

$$\sup \{ \|\xi^+(\omega)\|_{\mathfrak{R}} : \omega \in \Omega^+, \|\omega\|_H \leq 1 \} = \|\xi^\#\|_{L(\Omega^\#, x)}.$$

Если бы  $\Omega^+$  оказалась собственным линейным множеством в  $H$ , то конструкция, приведённая в первой части доказательства, позволила бы продолжить (элементарным расширением)  $\xi^+$  на большее линейное множество с сохранением всех нужных свойств, а это противоречило бы максимальнойности гнезда  $\Sigma^+$ . Поэтому  $\Omega^+ = H$  и, таким образом,  $\xi^+ \in \mathcal{L}(H, x)$  при  $\|\xi^+\|_{L(H, x)} = \|\xi^\#\|_{L(\Omega^\#, x)}$ . Доказательство завершено.

Условие (на структуру пространства  $x$ ) распространения линейного оператора с сохранением операторной нормы, сформулированное в предложении 1, является так же и необходимым (см. ниже предложение 2), для установления этого свойства нам понадобятся некоторые приготовления, скомпонованные в лемме 1.

Лемма 1.<sup>7</sup> Допустим, что  $S$  – произвольное бесконечное множество и  $l^\infty(S)$  – семейство всех заданных на  $S$  вещественных ограниченных функций. Тогда, если для любых фиксированных функций  $f', f'' \in l^\infty(S)$  положить

$$\|f', f''\|_\infty = \sup \{ |f'(s) - f''(s)| : s \in S \}, \tag{5}$$

то  $(l^\infty(S), \|\cdot\|_\infty)$  – банахово пространство типа  $\mathfrak{R}$ .

Доказательство аналогично рассуждениям, приведенным в сноске 3.

Предложение 2. Банахово пространство  $(x, \|\cdot\|_x)$  будет типа  $\mathfrak{R}$ , если какое бы банахово пространство  $(H, \|\cdot\|_H)$  ни взять, всякий непрерывный линейный оператор, отображающий произвольное подпространство из  $H$  в  $x$ , допускает линейное непрерывное распространение с сохранением операторной нормы на всё  $H$ .

Доказательство. Пусть  $S$  – шар сопряженного пространства  $x^*$ . Тогда, очевидно,  $x^{**}$  – подпространство в  $l^\infty(S)$ , при этом существует (см. [13, с. 241]) линейная изометрия пространства  $x$  на подпространство в  $x^{**}$ , таким образом, установили, что  $x$  суть подпространство в  $l^\infty(S)$ . Теперь рассмотрим тождественный оператор  $I_x: x \rightarrow x$ . По условию предложения оператор  $I_x$  допускает распространение  $I_L: l^\infty(S) \rightarrow x$ , такое, что операторная норма  $I_L$  совпадает с нормой оператора  $I_x$  и равна 1. Следовательно,  $I_L$  – линейный непрерывный проектор из  $l^\infty(S)$  на  $x$ , имеющий норму 1. Согласно леммы 1  $l^\infty(S)$  – пространство типа  $\mathfrak{R}$ . Очевидно, что тип  $\mathfrak{R}$  сохраняется при проектировании с нормой единица, поэтому  $x$  – пространство типа  $\mathfrak{R}$ . Предложение доказано.

Замечание 2. Установлено (см. [14]), что банахово пространство обладает типом  $\mathfrak{R}$  в том и только в том случае, если оно линейно изометрично банахову пространству  $C(Q)$ , где  $Q$  – экстремально несвязный компакт; напомним, компакт  $Q$  называется экстремально несвязным, если замыкание всякого открытого множества из  $Q$  открыто-замкнуто [13, с. 366] (так например, – любое

<sup>6</sup> Обычная математическая индукция доказывает, что теорема верна для всех конечных положительных целых чисел, упорядоченных «по возрастанию». Трансфинитная индукция по существу использует тот же метод, но распространяет его на вполне упорядоченные множества трансфинитных ординальных чисел.

<sup>7</sup> Не трудно показать, теорема I.2.19 [6, с. 35], что  $l^\infty(S)$  с метрикой  $\|\cdot\|_\infty$  – полное метрическое пространство, сходимость в котором означает равномерную сходимость соответствующей последовательности функций.



конечное множество с дискретной топологией; в разделах 3, 4 более детально рассмотрим этот вариант  $Q$ ).

Для описанной выше конструкции аналитического решения задачи расширения линейного непрерывного оператора, сохраняющего норму, существует двойственный процесс (предложение 3), который установил и развил Какутани [16].

**Предложение 3.** Пусть  $(H, \|\cdot\|_H)$  – фиксированное нормированное пространство,  $E$  – произвольное линейное множество в  $H$ ,  $X$  – произвольное банахово пространство и пусть  $\xi_-: E \rightarrow X$  – линейный непрерывный оператор. Тогда, чтобы существовал  $\xi^+$  – линейный непрерывный оператор такой, что

$$\xi^+: H \rightarrow X, \quad \xi^+ \supset \xi_-, \quad \|\xi^+\|_{L(H,X)} = \|\xi_-\|_{L(E,X)},$$

необходимо и достаточно, чтобы  $H$  по структуре было предгильбертово, т.е.

$$\|x + y\|_H^2 + \|x - y\|_H^2 = 2(\|x\|_H^2 + \|y\|_H^2), \quad \forall x, y \in H.$$

**Доказательство.** Ограничимся рассмотрением достаточности, поскольку именно к ней в дальнейшем (теорема 2) и будем апеллировать, при этом будем считать, что  $(H, \|\cdot\|_H)$  – гильбертово пространство. Обозначим через  $\hat{E}$  замыкание множества  $E$ ;  $\hat{E}$  – подпространство пространства  $H$ . Так как  $E$  плотно в  $\hat{E}$ , то распространение оператора  $\xi_-$  на  $\hat{E}$  с сохранением нормы возможно по теореме 2 [13, с. 245]. Поэтому, не уменьшая общности, можно с самого начала считать, что  $E$  – подпространство. Далее, обозначим через  $P$  оператор проектирования [13, с. 202] пространства  $H$  на подпространство  $E$  и положим  $\xi^+(x) = \xi_- \circ P(x)$ ,  $\forall x \in H$ . Тогда линейность, непрерывность и  $\|\xi^+\|_{L(H,X)} = \|\xi^{\#}\|_{L(E,X)}$  оператора  $\xi^+$  вытекает из

$$\begin{aligned} \|\xi^+(x)\|_X &\leq \|\xi_-\|_{L(E,X)} \|P(x)\|_E \leq \\ &\leq \|\xi_-\|_{L(E,X)} \|P\|_{L(H,E)} \|x\|_H = \|\xi_-\|_{L(E,X)} \|x\|_H, \end{aligned}$$

так как  $\|P\|_{L(H,E)} = 1$  согласно пункту d) [13, с. 202]. Доказательство завершено.

Предложения 1–3 составляют (как скоро убедимся) «краугольные камни» геометрического фундамента в решении задачи дифференциальной реализацией оптимальной по критерию операторной нормы  $\|\cdot\|_{L(H,X)}$ . Правда остается открытым вопрос:<sup>8</sup> как «согласуется» структура бесконечномерного сепарабельного банахова равномерного выпуклого пространства с геометрией пространства типа  $\mathfrak{R}$ ? Какой в такой постановке можно получить результат, показывает следующая лемма.

**Лемма 2.** Предположим, что  $(X, \|\cdot\|_X)$  – некоторое вещественное бесконечномерное сепарабельное банахово равномерно выпуклое пространство. В этом положении найдется такое множество  $S$ , что  $X$  – замкнутое подпространство пространства  $(l^\infty(S), \|\cdot\|_\infty)$ , при этом норма  $\|\cdot\|_X$  – сужение на  $X$  нормы  $\|\cdot\|_\infty$ .

**Замечание 3.** В разделах 3, 4 ограничимся рассмотрением локально-компактных  $l^\infty(S)$ , поэтому не будем в этом разделе глубже вдаваться в вопросы сепарабельности подпространств в  $l^\infty(S)$ .

**Доказательство.** Пусть  $S$  – замкнутый шар единичного радиуса с центром в нуле сопряженного пространства  $X^*$ . Далее, поскольку  $X$  будет рефлексивно (см. теорему 2 [7, с. 182]), то  $X$  – замкнутое (в силу полноты  $X$ ) подпространство в  $l^\infty(S)$ , при этом согласно пункту (b) теоремы 4.3 [17, с. 106] и формулы (5) будут выполнены следующие «сочлененные» равенства:

$$\|x\|_X = \sup \{ |\langle x, s \rangle_X| : s \in S \} = \|x\|_\infty,$$

где  $\langle \cdot, \cdot \rangle_X$  – каноническая билинейная форма, устанавливающая двойственность между пространствами  $X$  и  $X^*$ . Этим лемма доказана.

<sup>8</sup> Не трудно показать (§ 4.6 [13, с. 40]), что любое бесконечномерное пространство  $l^\infty(S)$  с метрикой  $\|\cdot\|_\infty$  всегда (!) несепарабельное.

Заметим, что построенный в приведенном доказательстве  $S$ -шар пространства  $X^*$  является  $\sigma(X^*, X)$ -компактным (теорема Алаоглу–Бурбаки [13, с. 118]), – иногда этот геометрический факт бывает полезен.

**3. Разрешимость задачи дифференциальной реализации с минимальной операторной нормой и локально-компактным пространством выходных сигналов**

Одна из целей, проводимых в этом разделе исследований, – выявить связь теории реализации и теории расширения операторов, которая обогащает каждый из этих разделов математического моделирования возможностью интерпретации в другом разделе. Такая взаимосвязь между различными разделами теории математического моделирования ломает зачастую искусственное её деление на различные предметные «области» и подчеркивает существенное единство этой теории.

Ниже понадобятся некоторые дополнительные конструкции. Обозначим через  $L(T, \mu, R)$  пространство классов  $\mu$ -эквивалентности всех вещественных  $\mu$ -измеримых на  $T$  функций и пусть  $\leq_L$  – квазиупорядочение в  $L(T, \mu, R)$  такое, что  $\phi_1 \leq_L \phi_2$ , когда  $\phi_1, \phi_2 \in L(T, \mu, R)$  и при этом  $\phi_1(t) \leq \phi_2(t)$   $\mu$ -почти всюду в  $T$ . Наименьшую верхнюю грань для подмножества  $W \in L(T, \mu, R)$  обозначим  $\sup_L W$ , если она существует для  $W$  в структуре частичного упорядочения  $\leq_L$ . Далее, примем  $U := X \times Y \times Z$  и введем *энтропийный оператор Релея–Ритца*  $\Psi: AC(T, X) \times L_q(T, \mu, Y) \times L_q(T, \mu, Z) \rightarrow L(T, \mu, R)$ , построенный согласно «поточечного правила» на промежутке  $T$  (для конечномерного  $U$  данная конструкция была впервые рассмотрена в статье [18]):

$$\Psi(g, w, v)(t) := \begin{cases} \|dg(t)/dt\|_X (\|g(t)\|_X^q + \|w(t)\|_Y^q + \|v(t)\|_Z^q)^{-1/q}, & \text{если } (g(t), w(t), v(t)) \neq 0 \in U, \\ 0 \in R, & \text{если } (g(t), w(t), v(t)) = 0 \in U; \end{cases}$$

корректность оператора  $\Psi$  определяется теоремой 2 [7, с. 182] и теоремой 2.1 [19, с. 16]; данные теоремы «обеспечивают» существование вектор-функции  $dg(\cdot)/dt$ . Везде далее  $\text{Fr}(\cdot)$  – граничный оператор [20, с. 51] на семействе подмножества некоторого (фиксированного) топологического пространства.

В общих предположениях раздела 1 (сепарабельность и равномерная выпуклость пространств  $X, Y, Z$ ) нельзя доказать общую теорему, дающую необходимые и достаточные условия формальной разрешимости дифференциальной реализации с минимальной операторной нормой. Поэтому, прежде всего, естественно поинтересоваться (в контексте определений 2, 3), когда все нормы в пространстве выходных сигналов  $(X, \|\cdot\|_X)$  эквивалентны  $(X, \|\cdot\|_{\mathfrak{R}})$ ? Например можно спросить какова при этом характеристика топологии в пространстве  $X$ , или (с учетом важного замечания 2) структура пространства  $C(Q)$ ; поскольку пространство  $X$  должно быть *нормируемо*, то оно как минимум должно быть *локально выпукло*, что ниже предполагается по умолчанию. Теорема 2 [13, с. 127] и теорема 1.22 [17, с. 24] содержат ответы на эти вопросы, – топология в  $X$  должна быть *локально компактна* (!) и это обеспечивает наличие (условие)  $\text{Card } Q < \infty$ ; важно отметить, что в такой постановке в геометрической структуре пространства входных сигналов  $H_q$  пространство  $(X, \|\cdot\|_X)$  во все не обязано быть типа  $\mathfrak{R}$ .

**Теорема 1.** Пусть  $N \in \Pi_{\#}$ ,  $\text{Sp}_N := \text{Span}\{(g, w, v) \in AC(T, X) \times L_q(T, \mu, Y) \times L_q(T, \mu, Z) : (g, w, v) \in N\}$  с топологией индуцированной из  $H_q$ ,  $\text{Fr}$  – граничный оператор в  $\text{Sp}_N$ ,  $\Theta$  – ограниченная выпуклая окрестность нуля в  $\text{Sp}_N$  и  $X$  – локально-компактно. Тогда пучок динамических процессов  $N$  обладает  $\xi_{\min}$ -моделью при  $(X, \|\cdot\|_{\mathfrak{R}})$  в том и только в том случае, если существует функция  $\phi \in L_p(T, \mu, R)$ ,  $0 \leq_L \phi$  такая, что

$$\sup_L \Psi(\text{Fr}(\Theta)) = \phi.$$

**Замечание 4.** Выпуклость окрестности  $\Theta$  не носит «принципиальный» характер (можно опустить), она лишь упрощает технические детали доказательства, исключая потенциально-возможные «патологии» геометрической структуры множества  $\Theta$  как «произвольной» окрестности нуля в линейном многообразии  $\text{Sp}_N$ .

**Доказательство** теоремы. Необходимость очевидна (теорема 2 [5]), докажем справедливость достаточных условий. С этой целью рассмотрим линейное множество  $\Omega^\# := \{\omega \in H_q; \exists T_r \mathcal{C}, \exists(g, w, v) \in \Theta, \omega = \chi_{T_r}(g, w, v)\}$ , где  $\chi_{T_r}$  – характеристическая функция подинтервала  $T_r = [t_0, t_r] \subset \mathcal{C}$ ,  $t_0 \leq t_r$ . Перейдем к частностям.

Условимся, что линейное множество  $\Omega^\#$  наделено топологической структурой, индуцированной из пространства  $H_q$ . В такой постановке, как показано в доказательстве теоремы 1 [21], формальный идентификационный процесс (I-процесс [22]), организованный на I-базисе  $\Omega^\#$ , определяет в силу теоремы 2 [23] единственную идентифицируемую  $\xi_p$ -модель в идентификационном пространстве  $(L(\Omega^\#, X), \mathfrak{S}_\Omega)$  (здесь  $\mathfrak{S}_\Omega$  – I-классификация на  $L(\Omega^\#, X)$ , имеющая I-базисе  $\Omega^\#$  [22, 23]), оператор которой обозначим через  $\xi^\#$ . Так как пространство выходных сигналов  $(X, \|\cdot\|_{\mathfrak{R}})$  является пространством типа  $\mathfrak{R}$ , то оператор  $\xi^\#$  в силу предложения 1 можно линейно с сохранением нормы распространить на  $H_q$ . Пусть  $\xi^+$  – такое распространение. Очевидно, что  $(A, B, B^\#)_p$ -модель  $\Gamma^{-1}(\xi^+)$  (её существование гарантировано леммой 2 [5]) будет  $(A, B, B^\#)_p$ -моделью в дифференциальной реализации семейства процессов  $N$ . Покажем, что оператор  $\xi^+$  отвечает прочим условиям  $\xi_{\min}$ -модели (по отношению к  $N$ ). Допустим, что в семействе  $(A, B, B^\#)_p$ -моделей для множества  $N$  существует такая, что для её  $\Gamma$ -образа  $\xi^*$  справедливо  $\|\xi^*\|_{L(H, X)} < \|\xi^+\|_{L(H, X)} = \|\xi^\#\|_{L(\Omega^\#, X)}$ . Но в этом случае будет  $\|\xi^*\|_{\Omega^\#} \|L(\Omega^\#, X)\| \leq \|\xi^*\|_{L(H, X)}$ ; тем самым пришли к противоречию, поскольку имеет место  $\xi^* | \Omega^\# = \xi^\#$ . Доказательство завершено.

К существованию  $\xi_{\min}$ -модели можно подойти с иной методологической позиции, нежели представленной теоремой 1. С этой целью в доказательстве заменим предложение 1 предложением 3, которое характеризует  $H_q$  в условиях, когда  $X$  – произвольное равномерно выпуклое пространство; данная характеристика определяет  $H_q$  при  $q = 2$  как *унитарное пространство* (когда  $X, Y, Z$  суть гильбертовы), поэтому ответ на такую методологическую постановку ниже даёт:

**Теорема 2.** Пусть  $N \in \Pi_\#$  и в  $L(H_2, X)$  пространство  $X$  равномерно выпукло (но не обязательно гильбертово). Тогда для дифференциальной реализации  $N$  существует  $\xi_{\min}$ -модель в  $L(H_2, X)$ , если  $H_2$  гильбертово (т.е.  $X, Y, Z$  – гильбертовы).

Как и следовало ожидать, вариант, когда пространство входных сигналов  $H_2$  гильбертово, а выходных сигналов  $X$  типа  $\mathfrak{R}$  обладает дополнительными преимуществами при построении  $\xi_{\min}$ -модели; присвоим такой  $\xi_{\min}$ -модели структуру  $H_2 \mathfrak{R}$  (или, равнозначно, –  $H_2 \mathfrak{R}$ - $\xi_{\min}$ -модель). Это особенно проявляется в том случае, когда множество  $N$  суть *одноэлементное*. Поэтому в оставшейся части будем считать, что  $\text{Card } N = 1$  (вполне частый практический случай) и

$$H_2 := L_2(T, \mu, R^n) \times L_2(T, \mu, R^m) \times L_2(T, \mu, R^k),$$

где  $R^n, R^m$  и  $R^k$  – евклидовы пространства, а  $R^n$  в  $L(H_2, R^n)$  имеет норму  $\|x\|_{\max, R^n} = \max\{|x_i|; 1 \leq i \leq n\}$ ,  $\text{col}(x_1, \dots, x_n) = x$  и  $N^* := \{(x^*, u^*, u^\#(x^*))\}$ , где  $(x^*, u^*, u^\#(x^*))$  – фиксированная тройка вектор-функций из  $\Pi_\# \subset H_2$ . При этом дополнительно предполагаем, что  $N^*$  имеет реализацию (1); все возможные затруднения снимаются, как только замечаем, что характеристика этого положения (в силу теоремы 1) является

$$\Psi(N^*) \in L_2(T, \mu, R),$$

где  $\Psi$  – оператор Релея–Ритца. Далее, условимся также, что  $\omega^* := \text{col}(x^*, u^*, u^\#(x^*))$  и, соответственно,  $\omega^*_{,i} := (\langle \omega^*, \omega^* \rangle_{R^{n+m+k}})^{-1} \text{col}(dx^*_1/dt, \dots, dx^*_n/dt)$ , где  $\langle \cdot, \cdot \rangle_{R^{n+m+k}}$  – операция скалярного произведения в пространстве  $R^{n+m+k}$ ,  $dx^*_i/dt$  –  $i$ -ый элемент вектора  $dx^*/dt$ . Доказательство следующей теоремы прозрачно в силу следствия 5 [24]<sup>9</sup>.

**Теорема 3.** Для  $N^*$  существует  $H_2 \mathfrak{R}$ - $\xi_{\min}$ -модель с  $\Gamma$ -изоморфной  $(A, B, B^\#)_2$ -моделью  $(A, B, B^\#) \in L_2$ , у которой аналитическое представление имеет вид

$$[A^*(t), B^*(t), B^{\#*}(t)] = \omega^*_{,i}(t) [\omega^*(t)]', \quad (6)$$

<sup>9</sup> Формула (6) – очевидная конкретизация общей конструкции (6) [24].

при этом норма  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели равна  $\max_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2}$ , где  $\omega_{di}^*$  –  $i$ -ая координата вектор-функции  $\omega_{di}^* [\cdot]'$  – операция транспонирования матрицы.

Теоремы 1–3 не отвечают на важный вопрос: имеет ли задача построения  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели единственное решение? Следующее утверждение даёт удивительно простой (и одновременно конструктивный) ответ на поставленный вопрос.

Теорема 4.  $\xi_{\min}$ -модель для реализации  $N^*$  в структуре  $H_2\mathfrak{R}$  единственна в том и только в том случае, если

$$\min_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2} = \max_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2}.$$

Пример 1. В предположении нарушенного условия ( $\min \dots = \max \dots$ ) теоремы 4, построим  $(A, B, B^\#)_2$ -модель,  $\Gamma$ -изоморфную некоторой  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели и не равную  $(A^*, B^*, B^{\#\#})$ . Пусть  $\min_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2} \neq \max_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2}$ . Тогда найдётся индекс  $j$ , для которого  $(\int_T \omega_{dj}^{*2}(t)\mu(dt))^{1/2} \neq \max_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2}$ .

Далее пусть число  $\alpha$  таково, что  $(\int_T \omega_{dj}^{*2}(t)\mu(dt))^{1/2} \leq \alpha < \max_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2}$ . Тогда любая вектор-функция  $s^*(\cdot) \notin H_2$  такая, что  $\text{supp}(\omega^*) \subset \text{supp}(s^*) \pmod{\mu}$  [13, с. 137],  $dx_j^*(t)/dt = \langle s^*(t), \omega^*(t) \rangle_{R^{n+m+k}} + k$  и  $(\int_T \langle s^*(t), s^*(t) \rangle_{R^{n+m+k}} \mu(dt))^{1/2} = \alpha$ , может заменить в качестве  $j$ -ой строки в функциональной (6) матрице  $\omega_{di}^* [\omega^*]'$  её  $j$ -тую (исходную) строку без потери у полученной таким образом  $(A, B, B^\#)_2$ -модели аналитического свойства, а именно, – быть  $\Gamma$ -изоморфной некоторой  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели множества  $N^*$ ; например, конструкция  $s^* = s_j^* + s$ , где  $s_j^*$  –  $j$ -ая строка матрицы  $\omega_{di}^* [\omega^*]'$  и  $s$  – любая вектор-функция из ортогонального дополнения –  $\Omega_\perp^*$  [13, с. 165] к максимальному  $I$ -базису  $\Omega^*$ , индуцированному процессом  $\omega^*$ , удовлетворяет всем выдвинутым к вектор-функции  $s^*$  требованиям.

Поскольку из  $\min_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2} \neq \max_{1 \leq i \leq n} (\int_T \omega_{di}^{*2}(t)\mu(dt))^{1/2}$   $(A, B, B^\#)_2$ -модель, осуществляющая «роль» реализации для наблюдаемого динамического процесса  $N^*$  с  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -моделью, не обладает свойством единственности, то процедуру её построения естественно усилить дополнительным требованием – о минимальности нормы этой  $(A, B, B^\#)_2$ -модели, как элемента банахова пространства  $L_2$ , что составит предмет рассмотрений следующего раздела.

#### 4. Конструирование $(A, B, B^\#)_2$ -моделей с минимальной нормой в $L_2(T, \mu, L(R^n, R^n)) \times L_2(T, \mu, L(R^m, R^m)) \times L_2(T, \mu, L(R^k, R^k))$

Операторная норма  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели для реализации  $N^*$  из § 3, в силу своей аналитической конструкции равна  $\max_{1 \leq i \leq n} (\int_T \langle s_i^*(t), s_i^*(t) \rangle_{R^{n+m+k}} \mu(dt))^{1/2}$ , где  $s_i^*$  –  $i$ -ая строка  $\Gamma$ -изоморфной для неё  $(A, B, B^\#)_2$ -модели. Таким образом, целесообразно рассматривать вектор-функции  $s_i^*$ ,  $1 \leq i \leq n$ , как элементы гильбертова пространства  $H_2 = L_2(T, \mu, R^n) \times L_2(T, \mu, R^m) \times L_2(T, \mu, R^k)$ . Сделаем этот факт «ключом» к решению задачи конструирования  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели в рамках Фурье-анализа.

Пусть для каждого индекса  $i=1, \dots, n$  выделены:  $G_i$  – бесконечномерное подпространство пространства  $H_2$ ,  $\{\varphi_{ij}\}_{j=1,2,\dots}$  – некоторая полная ортогональная система в  $G_i$  и  $s'_i \in G_i$  –  $i$ -ая объединённая (совместная) вектор-строка некоторой фиксированной нестационарной  $(A, B, B^\#)_2$ -модели  $(A', B', B^{\#\#}) \in L_2$ . Тогда, из  $s'_i \in G_i$ ,  $i=1, \dots, n$  следует, что для тройки матричных функций  $(A', B', B^{\#\#})$  корректно её аналитическое представление в структуре пространства  $L_2(T, \mu, L(R^{n+m+k}, R^n))$  (через ряды Фурье в гильбертовом пространстве  $H_2$ ) следующего вида:

$$[A', B', B^{\#\#}] = \begin{bmatrix} \sum_{j=1,2,\dots} c_{1j} \varphi_{1j} \\ \dots \\ \sum_{j=1,2,\dots} c_{nj} \varphi_{nj} \end{bmatrix}, \quad c_{ij} \in \mathbb{R}, \quad \varphi_{ij} \in G_i, \quad (7)$$

при этом [25, с. 411] для каждой вектор-строки  $s'_i(\cdot)$  ( $1 \leq i \leq n$ ) матрицы  $[A', B', B^{\#\#}]$  при  $q \rightarrow \infty$   $\mu$ -почти всюду на  $T$  будет выполняться поточечная сходимость:

$$\sum_{j=1,\dots,q} c_{ij} \varphi_{ij}(t) \xrightarrow{q \rightarrow \infty} s'_i(t). \quad (8)$$

В такой постановке упорядоченную тройку  $(\{G_i\}_{1 \leq i \leq n}, \{\varphi_{ij}\}_{1 \leq i \leq n, j=1,2,\dots}, (A', B', B^{\#\#}))$

будем называть<sup>10</sup> *сигнатурой «спектрально-векторной» идентификации* (сигнатурой СВИ) нестационарной  $(A, B, B^\#)_2$ -модели  $(A', B', B'^\#)$  динамического объекта (1). Отметим, что введение в рассмотрение процедуры СВИ методологически оправдано тем, что задача *конструирования* объекта (1), как правило, решается в условиях, когда *a priori* нет «прямых указаний» на вид аналитического представления (моделирования) нестационарных параметров  $(A, B, B^\#)_2$ -модели.

Основное свойство, характеризующее понятие сигнатуры СВИ как элемента формализации общей постановки задачи СВИ, свяжем с конструкцией семейства процессов  $P \in \Pi_{i\#}$ , имеющего реализацию (1) и порождающего I-базис, теоретико-множественная модель которого допускает конструктивное построение вида (7).

**Определение 4.** Пусть  $P \in \Pi_{i\#} \subset \mathcal{H}_2$ ,  $(\{G_i\}_{1 \leq i \leq n}, \{\Phi_{ij}\}_{1 \leq i \leq n, j=1,2,\dots}, (A', B', B'^\#))$  некоторое семейство динамических процессов и сигнатура СВИ. В этом случае будем говорить, что семейство  $P$  и сигнатура  $(\{G_i\}_{1 \leq i \leq n}, \{\Phi_{ij}\}_{1 \leq i \leq n, j=1,2,\dots}, (A', B', B'^\#))$  находятся в двойственности, если выполнены оба следующих условия<sup>11</sup>:

а) множество  $P$  удовлетворяет классу допустимых  $K$ -решений дифференциальной системы (1) с  $(A, B, B^\#)_2$ -моделью равной  $(A', B', B'^\#)$ ;

б) в пространстве входных сигналов  $H_2$  существует такой I-базис, индуцированный множеством  $P$ , что  $\forall i=1, \dots, n$  этот I-базис разделяет точки в  $G_i$ .

Через  $2^N$  обозначим семейство всех подмножеств фиксированного множества динамических процессов  $N \in \Pi_{i\#} \subset \mathcal{H}_2$ ; по существу элементы из  $2^N$  – «потенциальные кандидаты» для I-процесса, призванные *a posteriori* определить структурный «фенотип» третьей компоненты сигнатуры СВИ.

**Определение 5.** Задача СВИ  $(A, B, B^\#)_2$ -модели объекта (1) частично решена (аналогично, вполне решена) над множеством динамических процессов  $N \in \Pi_{i\#} \subset \mathcal{H}_2$ , если одновременно имеют место:

а) в  $2^N$  выделено непустое максимальное множество  $P$ , для которого существует реализация в виде некоторой дифференциальной системы (1) (соответственно показано, что  $N$  –  $K$ -решения некоторой дифференциальной системы (1));

б) определена некоторая сигнатура СВИ, двойственная множеству  $P$  из а) (аналогично, двойственная  $N$ ) такая, что в этой сигнатуре для тройки матричных функций  $(A', B', B'^\#)$  и любой пары индексов  $(i, j)$ ,  $1 \leq i \leq n, j=1,2,\dots$  возможно тем или иным способом вычисление коэффициента  $c_{ij}$  из разложения (7).

**Замечание 5.** Вводя в обращение понятия как *частичной* так и *вполне* разрешимости задачи спектрально-векторной идентификации, пунктом а) обеспечиваем её решение в классе  $(A, B, B^\#)_2$ -моделей, тогда как б) гарантирует единственность этого решения с «точностью» до компонент сигнатуры СВИ.

Определение 5 позволяет формально истолковывать задачу СВИ как некое «правило» или, если угодно, «закон соответствия», которое каждому заданному множеству наблюдаемых динамических процессов класса  $\Pi_{i\#} \subset \mathcal{H}_2$  сопоставляет предельные пары «*максимальное апостериорное семейство  $K$ -решений некоторой дифференциальной системы (1), двойственная сигнатура СВИ*», в терминах которых можно конструктивно построить разложение вида (7). Следует, однако, иметь в виду, что аналитические решения задачи СВИ, удовлетворяющие лишь правилам а) и б), могут иметь самую замысловатую структуру. Так, например, когда  $N$  – конечное множество, то допустим вариант, при котором задача СВИ вполне разрешима для любой линейной комбинации из  $N$ ,

<sup>10</sup> Здесь преднамеренно, ставим «...», так как «спектральная» терминология двусмысленна с учётом классических разделов теории операторов. Вместо неё уместнее говорить (употреблять) о «гармоническом» варианте термина. Однако необходимо признать, что «спектральные доводы» настолько укрепились в специальной литературе по ортогональным разложениям [26,27], что, по-видимому, бесполезно настаивать на изменении.

<sup>11</sup> Чтобы получить более выпуклое представление о возможном прикладном аспекте пунктов а) и б) определения 4, полезно (интересуясь «рецептурной» стороной вычислений) обратиться к рассмотренной в монографии [26] задаче *спектральной* идентификации элементов матриц  $A(\cdot)$ ,  $B(\cdot)$  и  $B^\#(\cdot)$  из дифференциальной системы (1) по замерам в точках интервала  $T$  вектор-функции  $(x(t), u(t), u^\#(x(t)))$ . Так как всё обсуждение в монографии [26] для данной задачи ведётся без анализа критерия, подобного пункту а), то даже в предположении, что алгоритмическая реализация полученного в [26] её (спектральной задачи) решения удовлетворяет пункту б), эта реализация не гарантирует обеспечения спектрально-параметрической идентифицируемости в том смысле, как это трактуется авторами указанной монографии, а именно, представление (вычисление) элементов матриц  $A(\cdot)$ ,  $B(\cdot)$  и  $B^\#(\cdot)$  в ортогональном базисе из  $L_2(T, \mu, R)$ .

тогда как над всем множеством  $N$  она разрешима лишь частично (см., например, конструкцию примера [24, с. 52]).

В случае, когда  $N$  – бесконечное множество, возможен вариант, при котором задача СВИ вполне разрешима для любого элемента из  $\text{Span } N$ , в то время как над всем множеством  $N$  задача СВИ не разрешима даже частично. В связи с этим полезно вводить те или иные дополнительные условия, выделяющие более узкие классы задач СВИ; вышеизложенное можно рассматривать как «идеологическую» подготовку к введению основного в этом разделе понятия, а именно,  $(A, B, B^\#)_{\min}$ -реализации, с помощью которой появляется (при апостериорном моделировании) возможность оперировать «оптимальными» сигнатурами СВИ.

Обозначим через  $(L^2, \|\cdot\|_{L^2})$  – пространство  $L_2$  с нормой (эквивалентной норме  $\|\cdot\|_L$  из § 1) вида  $\|(A, B, B^\#)\|_{L^2} := (\sum_{1 \leq i \leq n} \sum_{1 \leq j \leq n} (\int_T a_{ij}^2(t) \mu(dt)) + \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq m} (\int_T b_{ij}^2(t) \mu(dt)))^{1/2} + \sum_{1 \leq i \leq n} \sum_{1 \leq j \leq k} (\int_T b_{ij}^{\#2}(t) \mu(dt))$ ,  $(A, B, B^\#) = ([a_{ij}], [b_{ij}], [b_{ij}^\#]) \in L^2$ ; в силу теоремы 8 [25, с. 162] линейное  $(A, B, B^\#)_2$ -многообразие  $(L^2, \|\cdot\|_{L^2})$  гильбертово.

Определение 6. Пусть  $P \subset \Pi_{i^\#} \subset H_2$ . Скажем, что упорядоченная тройка матричных функций  $(A', B', B'^\#) \in L^2$  является  $(A, B, B^\#)_{\min}$ -реализацией над  $P$ , если имеют место оба следующих условия:

- $P$  есть семейство  $K$ -решений системы (1) с  $(A, B, B^\#)_2$ -моделью  $(A', B', B'^\#)$ ;
- упорядоченная тройка  $(A', B', B'^\#)$  имеет в  $(L^2, \|\cdot\|_{L^2})$  минимальную норму среди всех норм от  $(A, B, B^\#)_2$ -моделей, реализующих через (1) семейство процессов  $P$ .

Наша ближайшая цель – установить: если множество  $P = N^*$  (где  $N^*$  из теоремы 3), то  $(A, B, B^\#)_{\min}$ -реализация над  $P$  существует, единственна (как элемент линейного  $(A, B, B^\#)_2$ -многообразия  $(L^2, \|\cdot\|_{L^2})$ ) и решает единственным образом (в отличие от теоремы 3) задачу  $H_2 \mathcal{R}\text{-}\xi_{\min}$ -модели для наблюдаемого динамического процесса  $N^*$ . При этом последнее обстоятельство указывает на то, что предложенное понятие  $(A, B, B^\#)_{\min}$ -реализации возникает естественным образом, а исследование структуры сигнатуры СВИ, содержащей данную  $(A, B, B^\#)_{\min}$ -реализацию, и находящейся в двойственности к  $N^*$ , носит сугубо прагматический характер.

Пусть  $N^* = \{(x^*, u^*, u^\#(x^*))\}$ , где динамический процесс  $(x^*, u^*, u^\#(x^*))$  обладает реализацией (1). Обозначим через  $\Lambda_2^*$  пространство сигнальных функций, отвечающее  $(x^*, u^*, u^\#(x^*))$ , через  $\Omega_2^*$  – максимальный I-базис [22], индуцированный  $(x^*, u^*, u^\#(x^*))$  и  $\Lambda_2^*$ , при этом из теоремы 1 [22] семейство функций  $\Lambda_2^*$  имеет представление:

$$\Lambda_2^* = L_2(T, \nu, R), \nu = \int (\|x^*(t)\|_{R^n}^2 + \|u^*(t)\|_{R^m}^2 + \|u^\#(x^*(t))\|_{R^k}^2) \mu(dt),$$

где  $\|\cdot\|_{R^n}$ ,  $\|\cdot\|_{R^m}$  и  $\|\cdot\|_{R^k}$  евклидовы, в свою очередь,  $\Omega_2^* = \{\lambda \cdot (x^*, u^*, u^\#(x^*)) : \lambda \in \Lambda_2^*\}$ .

Теперь на  $L(H_2, R^n)$  рассмотрим [22, 28] формальный I-процесс<sup>12</sup>

$$\{H_2, R^n, (L(H_2, R^n), \cap \{\pi_\omega^{*-1}[\Delta_{R^n}] : \omega \in W\}, \cap \{\pi_\omega^{-1}[\xi(\omega)] : \omega \in W\}) : \exists W \subset \Omega_2^*\},$$

где  $\pi_\omega : L(H_2, R^n) \rightarrow R^n$  – отображение вычисления в точке  $\omega \in \Omega_2^*$  и  $\pi_\omega^*$  – его каноническое распространение [28],  $\xi(\omega) = \xi(\lambda \cdot (x^*, u^*, u^\#(x^*))) = \int_T \lambda(t) dx^*(t)/dt \mu(dt)$ ,  $\lambda \in \Lambda_2^*$  (см. формулу (3) [22]),  $\Delta_{R^n}$  – диагональ в  $R^n \times R^n$ . В силу теоремы 4 [22] данный I-процесс выделяет в пространстве  $L(H_2, R^n)$  тождественный (в соответствии с леммой 2 [5]) нетривиальный (не одноточечный) класс  $\xi_2$ -моделей, равный

$$\cap \{\pi_\omega^{-1}[\int_T \lambda(t) dx^*(t)/dt \mu(dt)] \subset L(H_2, R^n) : \omega = \lambda \cdot (x^*, u^*, u^\#(x^*)), \lambda \in \Lambda_2^*\}.$$

Этому классу на основании леммы 2 [5] в  $L^2$  отвечает класс  $(A, B, B^\#)_2$ -моделей, который, обозначим через  $L^*$ ; ясно, что любая система (1) с  $(A, B, B^\#)_2$ -моделью из  $L^*$  содержит в множестве своих допустимых троек типа « $K$ -решение, программное управление, позиционное управление» означенную выше тройку  $(x^*, u^*, u^\#(x^*))$ , причём класс  $L^*$  по данному признаку является максимальным в пространстве  $L^2$ .

Лемма 3. Класс  $L^*$  замкнут в  $L^2$ .

<sup>12</sup> В действительности необходимо показать корректность (определение 2 [28]) данного I-процесса, что при необходимости (если встать на формальную позицию) не составит труда ввиду «реализуемости»  $(x^*, u^*, u^\#(x^*))$ .

Принимая во внимание, что многообразие  $\mathbf{L}^*$  является классом смежности по некоторому подпространству пространства  $\mathbf{L}^2$ , и учитывая тот факт, что  $\mathbf{L}^2$  имеет структуру гильбертова пространства, можно подвести аналитический итог:

**Теорема 5.**  $(A, B, B^\#)_{\min}$ -реализация над  $N^*$  существует, единственна, принадлежит  $\mathbf{L}^*$ , решает задачу  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели над  $N^*$  и имеет вид (6).

Определим аналитическое представление сигнатуры СВИ, находящейся в двойственности к  $N^*$  и такой, что её структура в качестве третьей компоненты содержит  $(A, B, B^\#)_{\min}$ -реализацию над  $N^*$ . С этой целью выделим в гильбертовом пространстве  $\Lambda_2^*$  полную ортогональную систему функций  $\{\phi_j\}$  (ясно, что в терминах меры  $\mu$  система  $\{\phi_j\}$  ортогональная с весом  $(\|x^*(\cdot)\|_{R^{n^2}} + \|u^*(\cdot)\|_{R^{m^2}} + \|u^\#(x^*(\cdot))\|_{R^{k^2}})$ ). Для определённости рассмотрим тригонометрическую систему, которую обозначим как  $\{\phi_{ij}^*\}$ . Тогда  $\{\phi_{ij}^*\} = \{1, \sin(2\pi j\tau(\cdot)/\nu(T)), \cos(2\pi j\tau(\cdot)/\nu(T))\}$ :  $\tau(t) = \nu([t_0, t])$ ,  $t \in T = [t_0, t_1]$ ,  $j = 1, 2, \dots$ . Далее, рассмотрим сигнатуру  $(\{\Omega_{ij}^*\}_{1 \leq i \leq n}, \{\phi_{ij}^*\}_{1 \leq i \leq n, j=1, 2, \dots}, (A^*, B^*, B^{\#\#}))$ , которую для удобства обозначим через  $\text{SIG}^*$ .

**Теорема 6.** Сигнатура  $\text{SIG}^*$  находится в двойственности к  $N^*$ .

**Доказательство.** На основании определения 2 и теоремы 5 для доказательства теоремы 6 достаточно установить, что для любого  $i=1, \dots, n$  имеет место  $s_i^* \in \Omega_2^*$ , где  $s_i^*$  –  $i$ -ая совместная вектор-строка матричной тройки  $(A^*, B^*, B^{\#\#})$ ; то, что  $\Omega_2^*$  – подпространство в  $H_2$  – следствие теоремы 7 [23].

Пусть  $\mathbf{L}^0$  – подпространство в  $\mathbf{L}^2$ , все точки которого эквивалентны (лемма 2 [5]) точкам подпространства  $\cap \{\pi_{\omega}^{-1}[0] : \omega \in \Omega_2^*\}$  пространства  $L(H_2, R^n)$ . В соответствии с леммой 2 [5] имеет место следующее утверждение:

$$\Upsilon_1: ([s_i^\perp]) \in \mathbf{L}^0 \Leftrightarrow \forall \omega \in \Omega_2^*, \forall i=1, \dots, n, \langle s_i^\perp, \omega \rangle_{H_2} = 0,$$

где  $\langle \cdot, \cdot \rangle_{H_2}$  – каноническая билинейная форма (скалярное произведение) в  $H_2$ . Так как  $\mathbf{L}^*$  – класс смежности в  $\mathbf{L}^2$  по  $\mathbf{L}^0$ , то в силу  $\Upsilon_1$  справедливо предположение:

$$\Upsilon_2: (\forall i=1, \dots, n) (\exists s_i^* \in \Omega_2^*) (\forall [s_i] \in \mathbf{L}^*): s_i = s_i^\perp + s_i^*, s_i^\perp \perp \Omega_2^*;$$

здесь  $\perp$  – знак ортогональности в гильбертовом пространстве  $H_2$ . Далее, т.к.

$$\|([s_i])\|_{L^2} = (\sum_{1 \leq i \leq n} \int_T \langle s_i^\perp(t), s_i^\perp(t) \rangle_{R^{n+m+k}} \mu(dt) + \sum_{1 \leq i \leq n} \int_T \langle s_i^*(t), s_i^*(t) \rangle_{R^{n+m+k}} \mu(dt))^{1/2}$$

и  $\|([s_i^*])\|_{L^2} = \min\{\|([s_i])\|_{L^2} : [s_i] \in \mathbf{L}^*\}$ , то выполнение  $\min\{\|([s_i])\|_{L^2} : [s_i] \in \mathbf{L}^*\}$  при наличии утверждения  $\Upsilon_2$  будет в том и только в том случае, если для любого  $i=1, \dots, n$  будет  $s_i^\perp = 0$  и, поскольку  $[s_i] \in \mathbf{L}^*$ , то  $\forall i=1, \dots, n: s_i^* \in \Omega_2^*$ . Доказательство завершено.

Сравнительно простое устройство сигнатуры  $\text{SIG}^*$  приводит к вопросу: как для произвольной пары индексов  $(i, j)$ ,  $1 \leq i \leq n, j=1, 2, \dots$  определить численное значение коэффициента  $c_{ij}$  в формуле (7), отвечающее данной сигнатуре? Конструктивный ответ на этот вопрос означал бы «окончательную точку» в рассмотрении проблемы вполне разрешимости задачи СВИ для  $H_2\mathfrak{R}\text{-}\xi_{\min}$ -модели и  $(A, B, B^\#)_{\min}$ -реализации над динамическим процессом  $N^*$ . Ниже установим основные элементы процедуры вычислений коэффициентов Фурье из разложения (7), а также определённые виды и аналитические условия сходимости предела (8).

Пусть в терминах, принятых нами выше,  $(x^*, u^*, u^\#(x^*))$  и  $[s_i^*] = [A^*, B^*, B^{\#\#}]$  – аналитические конструкции, определяющие структуру сигнатуры  $\text{SIG}^*$ . Тогда

$$s_i^* = \lambda_i \cdot (x^*, u^*, u^\#(x^*)) \quad (1 \leq i \leq n), \quad (9)$$

при некоторой сигнальной функции  $\lambda_i \in \Lambda_2^*$ , своей (единственной по mod  $\nu$ ) для каждого индекса  $i$ ; всюду далее под обозначениями  $s_i^*$ ,  $\lambda_i$  ( $1 \leq i \leq n$ ) будем понимать конструкции из (9). В соответствии с формулой (9) правило определения коэффициентов Фурье в разложении (7) для случая  $N^*$  даёт следующее предположение.

**Теорема 7.** Пусть  $t \rightarrow \tau(t) := \nu([t_0, t])$ , тогда:

а)  $s^*_i(t) = (\alpha_{i0} + \sum_{j=1,2,\dots} (\alpha_{ij}\sin(2\pi j\tau(t)/\nu(T)) + \beta_{ij}\cos(2\pi j\tau(t)/\nu(T))))(x^*(t), u^*(t), u^\#(x^*(t)))$  с коэффициентами  $\alpha_{ij}, \beta_{ij}, (1 \leq i \leq n)$  равными

$$\begin{aligned} \alpha_{i0} &= (\int_T \lambda_i(t)\nu(dt))/\nu(T), \\ \alpha_{ij} &= (\int_T \lambda_i(t)\sin(2\pi j\tau(t)/\nu(T))\nu(dt))/\nu(T), \\ \beta_{ij} &= (\int_T \lambda_i(t)\cos(2\pi j\tau(t)/\nu(T))\nu(dt))/\nu(T), \end{aligned}$$

б)  $\|(A^*, B^*, B^{\#*})\|_{L^2} = (\sum_{1 \leq i \leq n} (\int_T \gamma_i^2(t)\nu(dt)))^{1/2}$ , где  $t \rightarrow \gamma_i(t) = (dx^*_i(t)/dt)(d\tau(t)/dt)^{-1}$  и  $\gamma_i = \lambda_i$ .

Замечание 6. Поскольку  $\text{supp}(dx^*_i/dt) \subset \text{supp}(d\tau/dt)$  по mod  $\mu$  (т.к.  $N^* - K$ -решение), то функции  $\gamma_i, 1 \leq i \leq n$  определены корректно.

Перейдём теперь к группе вопросов, связанных с рассмотрением свойства сходимости предела (8) при его «реализации» в структуре сигнатуры  $SIG^*$ .

Теорема 8. Для фиксированной точки  $t \in T$  предел (8) сходится к  $s^*_i(t) \in R^{n+m+k}$  в структуре  $SIG^*$ , если при некотором  $\delta > 0$  существует  $\int_{T_\delta} (\lambda_i(t+z) - \lambda_i(t))/|z|\nu(dz)$ , где  $T_\delta = [t-\delta, t+\delta] \cap T$ , а  $s^*_i$  и  $\lambda_i$  связаны соотношением (9).

Доказательство. Обозначим через  $s^q_i$  частичную сумму ряда Фурье для  $s^*_i$  в структуре  $SIG^*$ . Тогда вектор-функция  $s^q_i$  в силу теоремы 7 имеет вид:

$$s^q_i(t) = (\alpha_0 + \sum_{j=1,\dots,q} (\alpha_j \sin(2\pi j\tau(t)/\nu(T)) + \beta_j \cos(2\pi j\tau(t)/\nu(T))))(x^*(t), u^*(t), u^\#(x^*(t)))$$

с формулами для  $\alpha_j, \beta_j$ , представленными соотношениями:

$$\begin{aligned} \alpha_0 &= (\int_T \lambda_i(t)\nu(dt))/\nu(T), \\ \alpha_j &= (\int_T \lambda_i(t)\sin(2\pi j\tau(t)/\nu(T))\nu(dt))/\nu(T) \\ \beta_j &= (\int_T \lambda_i(t)\cos(2\pi j\tau(t)/\nu(T))\nu(dt))/\nu(T). \end{aligned}$$

Используя эти соотношения, а также формулу (9), определим в точке  $t \in T$  величину разности  $s^q_i(t) - s^*_i(t)$  в норме евклидова пространства  $R^{n+m+k}$ :

$$\begin{aligned} \|s^q_i(t) - s^*_i(t)\|_{R^{n+m+k}} &= |\int_T (\lambda_i(t+z) - \lambda_i(t)) \sin(2\pi(2q+1)\tau(t)/\nu(T)) / \sin(2\pi\tau(t)/\nu(T)) \nu(dz)| \cdot \\ &\cdot \|(x^*(t), u^*(t), u^\#(x^*(t)))\|_{R^{n+m+k}} / \nu(T). \end{aligned}$$

Таким образом, вопрос о сходимости предела  $s^q_i(t) (q \rightarrow \infty)$  из формулы (8) последовательно свели к вопросу о стремлении к нулю интеграла с ядром Дирихле при наличии условий Дини. Дальнейшие выкладки опускаем, отсылая за деталями (если они необходимы) к выводу теоремы 1 [25, с. 409].

Оценка взаимной близости элементов пространства  $C(T, R^{n+m+k})$  в топологии sup-нормы позволяет более тоньше (в сравнении с топологией поточечной сходимости) выяснить степень отклонения между частичной суммой ряда (8) и вектор-функцией  $s^*_i$ . Всё это даёт основание подробнее рассмотреть условия равномерной сходимости предела (8); ясно, что, если вектор-функция  $s^*_i$  имеет хотя бы один разрыв на  $T$ , то её ряд Фурье не может сходиться к ней равномерно и, таким образом, непрерывность отображения  $s^*_i: T \rightarrow R^{n+m+k}$  есть необходимое (но, конечно, не достаточное) условие равномерной сходимости предела (8).

Теорема 9. Пусть  $N^* \subset AC(T, R^n) \times C(T, R^m) \times C(T, R^k)$ . В этом случае предел (8) сходится к  $s^*_i \in C(T, R^{n+m+k})$  в сигнатуре  $SIG^*$  равномерно на  $T = [t_0, t_1]$ , если выполняется хотя бы одно из двух следующих условий:

а)  $(\forall \varepsilon > 0) (\exists \delta > 0) (\forall t \in T): \int_{T_\delta} (\lambda_i(t+z) - \lambda_i(t))/|z|\nu(dz)$ , где  $T_\delta = [t-\delta, t+\delta] \cap T$ ;

б)  $\lambda_i \in AC(T, R) \& d\lambda_i/dt \in L_2(T, \nu, R) \& \lambda_i(t_0) = \lambda_i(t_1)$ .

Схема доказательства теоремы 8, с учётом теорем 2 и 3 [25, с. 409], с модификацией (которую, в виду её прозрачности, опускаем) переносится на случай равномерной сходимости предела (8), иначе – на теорему 9.

Теорема 9 утверждает, что в некоторых случаях процедура аппроксимации нестационарных вектор-строк матричной модели  $(A, B, B^\#)_{\min}$ -реализации конструкциями частичных сумм векторных рядов Фурье вида  $-s^q_i(t), 1 \leq i \leq n$  приближает  $(q \rightarrow \infty)$  к «физически» реализуемой  $(A, B, B^\#)_2$ -модели класса  $C(T, L(R^n, R^n)) \times C(T, L(R^m, R^m)) \times C(T, R^k)$ ; правда это приближение ограничивается выполнением пунктов а) или б). Если «поискать» для варианта б) более слабые условия на конструкцию подсемейства сигнальных функций  $\{\lambda_i\}_{1 \leq i \leq n} \subset AC(T, R)$ , а именно,  $\{\lambda_i\}_{1 \leq i \leq n} \subset C(T, R)$ , то (привлекая теорему



Фейера [25, с. 415]) небольшое изменение в процедуре предела (8) приводит к такому же качественному результату, как и теорема 9, а именно, равномерной сходимости к вектор-функции  $s^*_i$ .

Теорема 10. Пусть  $N^* \in \mathcal{AC}(T, R^n) \times C(T, R^m) \times C(T, R^k)$ ,  $\lambda_i \in C(T, R)$ ,  $\lambda_i(t_0) = \lambda_i(t_1)$  и пусть  $s_{iq}(t) := q^{-1} s^q_i(t)$ , где  $s^q_i$  – частичная сумма ряда Фурье в сигнатуре  $SIG^*$ . Тогда сходимость  $s_{iq} \xrightarrow{q} s^*_i$  – равномерная сходимость на  $T = [t_0, t_1]$ .

(Ясно, что имеем:  $N^* \in \mathcal{AC}(T, R^n) \times C(T, R^m) \times C(T, R^k) \& [\text{supp } N^*]_T = T \Rightarrow \lambda_i \in C(T, R)$ ,  $(1 \leq i \leq n)$ , где  $[\text{supp } N^*]_T$  – замыкание носителя  $\text{supp } N^*$  в интервале  $T$ ).

### Резюме и комментарии

Норма оператора является мерой того, насколько он «усиливает» входные сигналы, измеряемые в соответствующей норме. Обычно желательно выбрать модель так, чтобы эта норма была по возможности мала – это означает, что выход системы будет мал при любых возмущениях, ограниченных в соответствующей норме. В этой парадигме в работе формализованы математические понятия (категории словаря математического моделирования динамики изучаемых процессов), такие как  $\xi_{\min}$ -модель,  $(A, B, B^\#)_{\min}$ -реализация,  $H_2 \mathcal{R} - \xi_{\min}$ -модель, процедура СВИ, позволяющие с общих позиций подходить к исследованию свойств существования и моделирования дифференциальных  $(A, B, B^\#)_2$ -моделей (описывающих «оптимальную» математическую структуру динамики изучаемых процессов), при этом формулировать утверждения о таких свойствах с помощью единой системы понятий и терминов, которая обеспечивает необходимую общность исследований, позволяя в то же время наиболее естественно излагать решения конкретных прикладных задач. При этом авторы ставили перед собой задачу подробно изложить применяемые конструкции и представить детальные доказательства всех сколько-нибудь существенных фактов; существуют, вообще говоря, два вида математических доказательств – экзистенциальные и конструктивные, выше были использованы примеры обоих типов. К общезначимым результатам теории структурно-параметрической идентификации, установленным в данной работе, относятся:

- определены необходимые и достаточные условия существования реализаций нестационарных  $(A, B, B^\#)_p$ -моделей с минимальной операторной нормой в зависимости от метрических структур пространств входных и выходных сигналов;

- в рамках Фурье-анализа обоснована и построена процедура спектрально-векторной идентификации матричного представления  $(A, B, B^\#)_2$ -модели в дифференциальной реализации динамического процесса  $N^* = \{x^*, u^*, y^*(x^*)\}$ .

Основная мотивация заключалась в рассмотрении на строгой основе (по существу теории Нахбина–Какутани) методологии, вовлечённой в этап *конструирования* на базе спектральной идентификации вектор-функций строк  $(A, B, B^\#)_2$ -модели  $\Gamma$ -изоморфной  $H_2 \mathcal{R} - \xi_{\min}$ -модели, и по-видимому, необходимо констатировать, что рассмотренная постановка задачи СВИ, а именно, для  $(A, B, B^\#)_{\min}$ -реализация для  $N^*$  носит, отчасти, паллиативный характер, поскольку не гарантирует восстановление спектральных характеристик вектор-строк системы (1) при  $\text{Card } N > 1$ . Данное затруднение перестаёт казаться столь принципиальным после того, как на основании модификации теоремы 4 [29] замечаем, что его можно обойти, если в постановке о вполне разрешимости задачи СВИ перейти (в отличие от варианта  $-N^*$ ) к более общему случаю, когда множество наблюдаемых динамических процессов  $N$  состоит не менее чем из  $n+m+k$  пар «траектория, программное управление, позиционное управление» (ясно, что данное условие необходимое, но не достаточное). Наконец заметим, что, если множество  $N$  таково, что задача СВИ над  $N$  не является вполне разрешимой, то проблему СВИ методологически можно ставить как задачу идентификации динамического объекта с наблюдателем.

Не менее содержательные вопросы в исследовании решения задачи СВИ в постановке  $\text{Card } N \geq n+m+k$ , по-видимому, возникнут и в отношении позиции б) определения 5; ясно, что выполнение именно этого условия сделает возможным представление *a posteriori* рядами Фурье вектор-строк в матричных  $(A, B, B^\#)_2$ -моделях в структуре фиксированных сигнатур СВИ (здесь необходимо отметить, что окончательное описание сигнатур СВИ такого рода, в частности, на основе быстрого преобразования Фурье [26,27], ещё ждёт своего окончательного решения). Что же касается моти-

вазии выбора структуры сигнатуры СВИ, отвечающей разложению именно вида (7)<sup>13</sup>, то заметим, что в действительности в переформулировке пунктов б) определений 4 и 5 в терминах матриц или их элементов (т.е. спектрально-матричной, т.к.  $L^2$  гильбертово, или спектрально-параметрической, т.к.  $L_2(T, \mu, R)$  гильбертово), а не вектор-строк  $(A, B, B^\#)_2$ -моделей, не вызывает каких-либо принципиальных трудностей (это – общий факт [7]), к тому же стоит иметь в виду, что некоторые важные вопросы (вычисление нормы  $(A, B, B^\#)_{\min}$ -реализации, анализ сходимости (8) и т.д.), как показывают теоремы 7–9, могут быть удовлетворительно решены уже на этой степени общности.

Подводя итог, отметим, что использование в определении 5 двух различных понятий а) и б) позволяет методологически разделить исследование задачи СВИ на две относительно независимые части, т.к. а) отражает геометрическую сторону её решения (и по существу носит характер *структурного тестирования*), б) – алгебраическую (выявляя аспект *параметрической идентификации*), при этом важность введения этих понятий состоит ещё и в том, что они порождают некоторый математический аппарат в моделировании динамики систем, расширяя проблематику и подсказывая новые пути и подходы, значение которых не ограничивается спецификой задачи СВИ.

### Литература

1. Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. М.: Мир, 1971. 400 с.
2. Willems J.C. System Theoretic Models for the Analysis of Physical Systems // Ric. Aut. 1979. № 10. P. 71–106.
3. Polderman J.W., Willems J.C. Introduction to mathematical systems theory: A behavioral approach. Berlin. Springer-Verlag, 1998. 454 p.
4. Русанов В.А., Лакеев А.В., Линке Ю.Э. Характеристический признак реализации нестационарной дифференциальной системы в банаховом пространстве // Доклады РАН. 2011. Т. 438, № 3. С. 323–325.
5. Русанов В.А., Антонова Л.В., Данеев В.А. К обратным задачам нелинейного системного анализа. Бихевиористический подход // Проблемы управления. 2011. № 5. С. 14–21.
6. Варга Дж. Оптимальное управление дифференциальными и функциональными уравнениями. М.: Наука, 1977. 624 с.
7. Иосида К. Функциональный анализ. М.: Мир, 1967. 624 с.
8. Сергиенко И.В., Дейнека В.С. Идентификация параметров эллипτικο-псевдопараболических распределенных систем // Кибернетика и системный анализ. 2011. № 4. С. 28–50.
9. Мамфорд Д. Лекции о кривых на алгебраической поверхности. М.: Мир, 1968. 236 с.
10. Еругин Н.П. Построение всего множества систем дифференциальных уравнений, имеющих заданную интегральную кривую // Прикладная математика и механика. 1952. Т. XVI, № 6. С. 659–670.
11. Месарович М., Такахара Я. Общая теория систем: математические основы. М.: Мир, 1978. 312 с.
12. Русанов В.А., Шарпинский Д.Ю. К теории структурной идентификации нелинейных многомерных систем // Прикладная математика и механика. 2010. Т. 74, вып. 1. С. 119–132.
13. Канторович Л.В., Акилов Г.П. Функциональный анализ. М.: Наука, 1977. 742 с.
14. Nachbin L.A. Theorem of the Hahn-Banach Type for Linear Transformations // Trans. Amer. Math. Soc. Vol. 68. № 1. 1950. P. 29–46.
15. Келли Дж. Л. Общая топология. М.: Наука, 1968. 384 с.
16. Kakutani S. Some characterization of Euclidean space // Jap. Journ. Math. 1939. V. 16. № 2. P. 93–98.
17. Рудин У. Функциональный анализ. М.: Мир, 1975. 448 с.
18. Данеев А.В., Русанов В.А., Шарпинский Д.Ю. Принцип максимума энтропии в структурной идентификации динамических систем. Аналитический подход // Известия вузов. Математика. 2005. № 11. С. 16–24.
19. Barbu V. Nonlinear Semigroups and Differential Equations in Banach Space. Noordhoff International Publishing, Leyden (the Netherlands), 1976. 352 p.
20. Энгелькинг Р. Общая топология. М.: Наука, 1986. 752 с.
21. Русанов В.А., Антонова Л.В. Геометрия пучков управляемых динамических процессов, обладающих нелинейной дифференциальной реализацией в равномерно выпуклом банаховом пространстве // Вестник Бурятского государственного университета. Вып. 9: Математика и информатика. 2011. С. 188–201.
22. Данеев А.В., Русанов В.А. К методам качественной теории идентификации сложных динамических систем // Доклады РАН. 1997. Т. 355, № 2. С. 174–177.
23. Данеев А.В., Русанов В.А. Геометрический подход к решению некоторых обратных задач системного анализа // Известия вузов. Математика. 2001. № 10. С. 18–27.

<sup>13</sup> Эта форма характеризуется своей лаконичностью и емкостью, открывая широкие возможности для анализа и синтеза соответствующих  $(A, B, B^\#)_2$ -моделей.

24. Данеев А.В., Лакеев А.В., Русанов В.А. К теории реализации сильных дифференциальных моделей. II // Сибирский журнал индустриальной математики. 2005. Т. VIII, № 2. С. 46–56.

25. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Наука, 1976. 544 с.

26. Солодовников В.В., Дмитриев А.Н., Егупов Н.Д. Спектральные методы расчета и проектирования систем управления. М.: Машиностроение, 1986. 440 с.

27. Залманзон Л.А. Преобразование Фурье, Уолша, Хаара и их применение в управлении, связи и других областях. М.: Наука, 1989. 496 с.

28. Данеев А.В., Русанов В.А. К аксиоматической теории идентификации динамических систем. I // Автоматика и телемеханика. 1994. № 8. С. 126–136.

29. Данеев А.В., Русанов В.А. К аксиоматической теории идентификации динамических систем. II // Автоматика и телемеханика. 1994. № 9. С. 120–133.

*Русанов Вячеслав Анатольевич*, д-р физ.-мат. наук, главный научный сотрудник, Институт динамики систем и теории управления СО РАН (ИДСТУ СО РАН); 664033, Россия, г. Иркутск, ул. Лермонтова, 134, ИДСТУ СО РАН; e-mail: [V.Rusanov@mail.ru](mailto:V.Rusanov@mail.ru); телефон: (3952) 36-50-93; факс: (3952) 51-16-16

*Антонова Лариса Васильевна*, канд. физ.-мат. наук, доц., Бурятский государственный университет, г. Улан-Удэ, e-mail: [antonov\\_vi@mail.ru](mailto:antonov_vi@mail.ru)

*Данеев Алексей Васильевич*, д-р техн. наук, проф., Иркутский государственный университет путей сообщения, e-mail: [daneev@mail.ru](mailto:daneev@mail.ru)

*Миронов Артем Сергеевич*, аспирант, Иркутский государственный университет путей сообщения, e-mail: [mironov\\_as@irgups.ru](mailto:mironov_as@irgups.ru)

*Rusanov Vyacheslav Anatolievich*, doctor of physical and mathematical sciences, principal researcher, Institute of system dynamics and control theory of Siberian branch of Russian academy of sciences – ISDCT SB RAS

*Antonova Larisa Vasilyevna*, candidate of physical and mathematical sciences, associate professor of Buryat state university

*Daneev Aleksey Vasilyevich*, doctor of technical sciences, professor of Irkutsk state transport university

*Mironov Artem Sergeevich*, postgraduate student of Irkutsk state transport university

## 4. Математическое моделирование

УДК 004.032.26(06)

©С.В. Архипов, А.Ф. Брагин

### НЕЙРОСЕТЕВОЙ ПОДХОД В ПОСТРОЕНИИ АДАПТИВНЫХ СЕТОК ДЛЯ НЕРЕГУЛЯРНЫХ ПЛОСКИХ ОБЛАСТЕЙ

В работе предлагается модифицированный алгоритм построения адаптивных сеток на сложных нерегулярных плоских областях, допускающих наличие разветвляющихся узких трещин, многосвязность, вырезы разных размеров и т.п. Осуществлена программная реализация метода, с возможностью интерактивного изменения параметров обучения нейросети. Улучшенная адаптация сеток на нерегулярных областях найдет свое применение в задачах численного моделирования физических процессов, обработке изображений, визуализации данных, в графических приложениях и т.д.

**Ключевые слова:** нейронные сети, самоорганизующиеся карты признаков, адаптивные сетки, дискретизация физической области.

*S.V. Arkhipov, A.F. Bragin*

### THE NEURAL NETWORK APPROACH IN CONSTRUCTION OF ADAPTIVE GRIDS FOR IRREGULAR FLAT AREAS

The paper introduces the modified algorithm of construction of adaptive grids on complex irregular flat areas which can have branching narrow cracks, multiple connection, cuts of different sizes and etc. Program realization of this method is carried out and the program interface allows an interactive variation of parameters of training neural network. The improved adaptation of networks on irregular areas will be used in problems of numerical modeling of physical processes, processing of images, visualization of data, graphic appendices and etc.

**Keywords:** neural networks, self-organizing cards of signs, adaptive grids, discretization of physical area.

#### Введение

Высокий интерес к проблеме построения адаптивных сеток обусловлен необходимостью дискретизации пространственной области для получения достоверного численного решения прикладных физических задач. Кроме того, адаптивные сетки широко используются при обработке изображений, визуализации данных, в графических приложениях, задачах кластеризации и т.д.

Основная идея методов построения адаптивных сеток состоит в уменьшении размеров ячеек в зонах расчетной области, соответствующих значительным ошибкам решения. К традиционным методам построения адаптивных сеток с заданной плотностью распределения можно отнести метод эквираспределения [1], Томпсона [2], эллиптический метод [3], алгебраические методы [4], конформных отображений [5] и т.д. Для получения качественных адаптивных сеток подобные методы требуют решения сложных систем нелинейных дифференциальных уравнений с частными производными, что накладывает ряд известных ограничений.

Сегодня высокую эффективность при построении адаптивных сеток с заданной плотностью на сложной физической области демонстрируют нейросетевые алгоритмы [7]-[10] на основе классической теории самоорганизующихся карт Кохонена (сеть SOFM - Self-Organizing Feature Maps, T.Kohonen) (например, [11], [12]).

В работе [9] показано, что применение базовой модели SOFM приводит к появлению граничного эффекта, наличию мертвых нейронов и нарушению гладкости сетки. Для решения указанных проблем предложены модифицированные методы, в основе которых лежит идея чередования базового алгоритма для внутренних и внешних узлов [7], использования так называемых раскрашенных моделей и специальных алгоритмов сглаживания [9]. Общеизвестный факт необходимости проведения многочисленных вычислительных экспериментов по выбору параметров обучения нейросети для приемлемой адаптации сетки инспирировал работы [13],[14], в которых, предложены рекомендации по выбору параметров функции расстояния и скорости обучения при адаптации плоской области регулярной сеткой модифицированным алгоритмом SOFM.

Однако, применение известных модифицированных нейросетевых алгоритмов к построению адаптивных сеток на сложных нерегулярных областях, характеризующихся многосвязностью, наличием разветвляющихся узких трещин, вырезов сложной формы, привело к наличию артефактов, самопересечений сети, а также выходов узлов и ребер сети за границы области.

В данной работе, предложен улучшенный композиционный [9] алгоритм, отличающийся пересчетом сеточного расстояния после удаления узлов вышедших за границы нерегулярной области. Осуществлена программная реализация метода, с возможностью интерактивного изменения параметров обучения нейросети. Проведены многочисленные вычислительные эксперименты подтверждающие эффективность разработанных алгоритмов на сложных нерегулярных областях.

### 1. Краткое описание нейросетевых алгоритмов

Работа сети SOFM характеризуется этапом инициализации карты и циклом:

1. выбор случайного образца  $x(n)$  с заданной плотностью распределения;
2. нахождение узла победителя (best matching unit, BMU) – кластера на карте признаков, вес которого имеет меньшее отличие в заданной метрике от случайного образца;
3. корректировка узлов из числа близлежащих к победителю – изменение веса победителя и его соседей с целью приближения к случайному образцу;
4. определение ошибки карты.

При определении узла победителя в шаге 2, как правило, в качестве метрики выбирается Евклидово расстояние  $d$ .

Корректировка положений узлов шага 3 происходит в зависимости от степени близости к победителю с помощью функции соседства  $\Theta(n, i_{BMU}, j_{BMU}, i, j)$  по формуле:

$$w_{ij}^{n+1} = w_{ij}^n + q(n, i_{BMU}, j_{BMU}, i, j) \cdot x(n) - w_{ij}^n, \quad (1)$$

где  $n$  – номер итерации,  $w_{ij}$  – вес  $ij$ -го узла,  $x(n)$  – случайно выбранный образец,  $i_{BMU}, j_{BMU}$  – индекс узла победителя для образца  $x(n)$ .

Функция соседства представляет собой невозрастающую функцию от дискретного времени  $n$  и расстояния между нейроном-победителем и соседними нейронами в сетке. Эта функция разбивается на две части: функцию расстояния  $h(d, n)$  и скорости обучения  $\delta(n)$  т.е.  $q(n, i_{BMU}, j_{BMU}, i, j) = \delta(n) \cdot h(d, n)$ .

Обычно применяется функция Гаусса

$$h(d, n) = e^{-\frac{d^2}{2\sigma^2(n)}}. \quad (2)$$

Функция  $\sigma(n)$  называется радиусом обучения, который выбирается достаточно большим на начальном этапе обучения и постепенно уменьшается так, что в конечном итоге обучается один нейрон-победитель. В качестве радиуса обучения используют линейно или экспоненциально убывающую функцию от времени, например, в работе [8]

$$\sigma(n) = a \cdot n^{-0.2}, \quad (3)$$

где величина  $a$  выбирается таким образом, чтобы на первой итерации получили осязаемое смещение все узлы карты. При этом в качестве расстояния  $d$  в формуле (2) предлагается использование сеточного расстояния

$$d^2 = (i_{BMU} - i)^2 + (j_{BMU} - j)^2. \quad (4)$$

Функция скорости обучения  $\delta(n)$  также представляет собой функцию, убывающую от времени

$$\delta(n) = n^{-0.2}. \quad (5)$$

Применение выше описанного алгоритма, который назовем базовым, в сочетании с различными вариантами функции соседства приводит к трем основным проблемам [9]:

1. Адаптация сетки на невыпуклой области  $G$  не гарантирует, что все узлы сетки будут принадлежать области  $G$ .
2. Граничные узлы построенной сетки расположены на определенном расстоянии до границы области, отличным от нуля. Это расстояние сопоставимо со средним расстоянием между узлами сетки.
3. Нарушение гладкости адаптивной сетки, вследствие уменьшения радиуса обучения на стадии уточнения.

Для решения двух первых проблем в работе [7] предложена идея модификации алгоритма обучения SOFM, состоящая в том, что чередуется применение этого алгоритма отдельно для граничных и внутренних узлов. Один цикл такого чередования называется макроитерацией [8]. В соот-

ветствии с этой идеей был разработан модифицированный алгоритм построения конечно-разностных адаптивных сеток [9].

Известный в литературе модифицированный алгоритм состоит из следующих основных этапов.

1. Инициализация положений узлов сетки.
2. На первой макроитерации ( $s=1$ ) применяется базовый алгоритм в течение  $n_0$  итераций ко всем узлам сетки.
3. На каждой макроитерации с номером  $s > 1$  выполняются следующие действия:
  - а) Применение базового алгоритма в течение  $n_1(s)$  итераций к граничным узлам сетки с генерацией точки только на границе области.
  - б) Применение базового алгоритма в течение  $n_2(s)$  итераций ко всем узлам с генерацией точки во всей области. При этом все граничные узлы зафиксированы и не меняют своего положения. Кроме того, если узлом победителем является граничный узел сетки, то он заменяет случайную точку  $x(n)$ .
4. Повторяются макроитерации до тех пор, пока изменения положений узлов не станут достаточно малыми.

## 2. Постановка задачи

Приведенный модифицированный алгоритм показывает высокую эффективность [9],[13],[14] на односвязных областях, без наличия узких трещин и других выраженных особенностей (рис. 1).

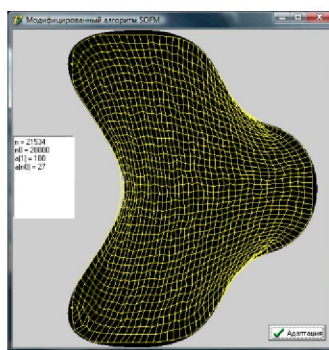


Рис. 1. Результат построения адаптивной сетки с применением модифицированного алгоритма на простой области

На рисунке 2 показан результат работы модифицированного алгоритма на нерегулярной области, отличающейся множеством узких вырезов. В реальной физической задаче такие вырезы могут соответствовать трещинам на поверхности твёрдого тела.

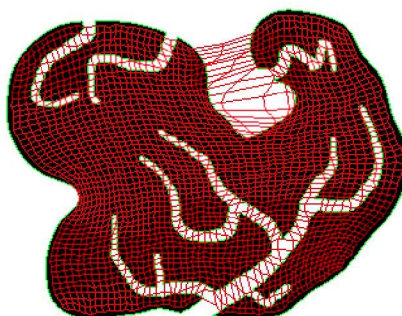


Рис. 2. Результат построения адаптивной сетки с применением модифицированного нейросетевого алгоритма на нерегулярной сложной области

Под нерегулярными областями в данной работе подразумеваются, прежде всего, сложные невыпуклые области с глубокими вырезами разной формы, области с множеством разветвляющихся трещин, многосвязные области и т.п. Примеры таких областей представлены на рисунке 3.



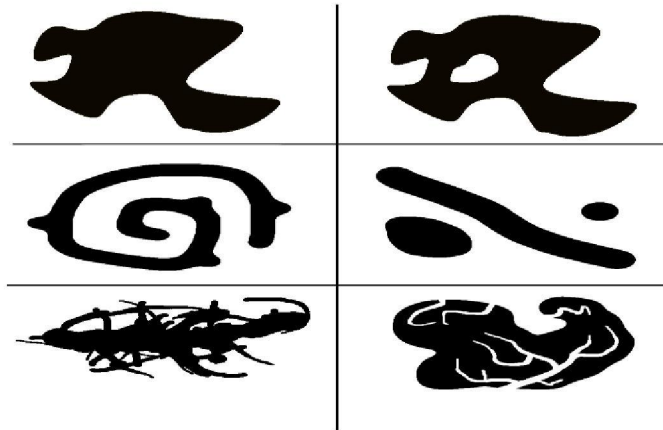


Рис. 3. Примеры сложных нерегулярных плоских областей

Для достижения улучшенной адаптации на многосвязных областях, областях с вырезами осуществляют [9] удаление нейронов, вышедших за границу области. Затем применяют описанный выше модифицированный алгоритм с учетом адаптации внешних и внутренних граничных узлов сетки к соответствующим границам области.

Такой подход в случаях узких вырезов, когда радиус обучения соизмерим с размерами выреза, не дает качественной адаптации границ сетки к границам области. В процессе выполнения второго уточняющего этапа обучения возникают самопересечения сети, артефакты.

Достаточно очевидным является предположение, что при ослаблении латеральных связей между узлами сети, лежащими на противоположных сторонах выреза в адаптируемой области, можно снизить указанный негативный эффект. Таким образом, для решения проблемы качественного построения адаптивных сеток на сложных нерегулярных областях необходимо

1. разработать способ ослабления латеральных связей между узлами сети, лежащими на противоположных сторонах выреза;
2. разработать критерий для оценки негативного эффекта выхода нейронов за границы области;
3. осуществить программную реализацию и апробацию разработанных методик.

### 3. Основные идеи

#### 3.1. Корректировка латеральных связей между узлами сети.

В модифицированном методе после первой макроитерации добавим этап удаления узлов вышедших за границу области. Причем удаление нейрона относительно начальной топологии сети влечет удаление смежных квадратов. В измененной таким образом сети уже не естественно использовать формулу сеточного расстояния (4), т.к. она не представляет собой функцию кратчайшего расстояния (рис. 4,а).

Изменим функцию  $d$  из (4), так чтобы она представляла собой функцию кратчайшего расстояния между двумя узлами сети на многосвязной области (рис. 4,б).

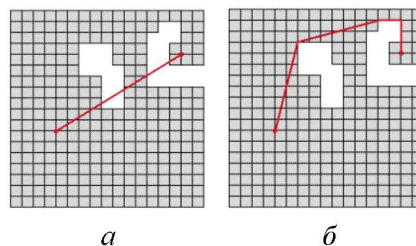


Рис. 4. Схема расчета сеточного расстояния  
 $a$  – без учета топологии сети,  $b$  – с учетом топологии сети

В отличие от односвязной области, определение расстояния между точками многосвязной области не является тривиальной задачей и требует отдельного решения.

В данном случае тот факт, что количество узлов сети конечно позволяет поставить эту задачу в терминах теории графов. В такой интерпретации узел сети – это вершина графа. Если для пары узлов существует соединяющий отрезок, лежащий в пределах многосвязной области, соответствующей топологии нейросети, то в графе существует ребро, вес которого равен длине отрезка.

Нужно отметить, что построенный таким образом граф, при умеренном количестве вырезов, является плотным.

Таким образом, функция  $d$  есть кратчайшее расстояние в графе между вершинами соответствующими узлу победителю и узлу, выбранному в процессе пересчета сеточного расстояния.

В данной работе для расчета кратчайшего расстояния выбран алгоритм Флойда-Уоршела. Это динамический алгоритм для нахождения кратчайших расстояний между всеми вершинами взвешенного ориентированного графа со сложностью  $O(n^3)$ .

Для построения графа, необходимо знать лежит ли отрезок соединяющий узлы нейросети в области определяемой топологией нейросети. Для односвязной области справедливо, что ей полностью принадлежат подобные отрезки. Многосвязная область нейросети модифицированного метода получается из односвязной вырезанием квадратных областей с центром в удалённом нейроне. Таким образом, вырезанные квадратные области являются единственным возможным препятствием между узлами сети. Следовательно, достаточно проверить пересекает ли отрезок некоторый квадрат, и положительный ответ будет говорить о том, что отрезок полностью не принадлежит области.

Для проверки пересекает ли отрезок квадрат, проверим на пересечение каждую его сторону. Для этого воспользуемся алгоритмом, суть которого заключается в последовательной проверке расположения вершин каждого отрезка относительно прямой содержащей другой отрезок.

Можно заметить, что подход к определению принадлежности отрезка области является затратным, т.к. возможна проверка лишних вырезанных областей, которые полностью перекрываются другими областями. Например, можно было бы объединить все вырезанные области и в некоторых случаях уменьшить количество проверяемых на пересечение отрезков.

Тем не менее, представленный алгоритм работоспособен и эффективен для данной задачи. Покажем отсутствие существенного проигрыша при его использовании. Количество вырезанных нейронов  $k$  не превышает количество всех нейронов  $n$ . Каждый вырезанный нейрон добавляет к списку проверяемых на пересечение отрезков ровно 4 отрезка. Асимптотическая сложность алгоритма проверки пересечения равна  $O(1)$ . Следовательно, сложность алгоритма проверки принадлежности отрезка области в целом равна  $4kO(1) = O(k)$ .

Алгоритм используется для подготовки графа в алгоритме Флойда-Уоршелла сложности  $O(n^3)$ . Тогда сложность итогового алгоритма пересчета сеточного расстояния для нейросети равна  $O(k) + O(n^3)$ , отсюда при  $k \leq n$  суммарная сложность алгоритма составит  $O(n^3)$ .

### 3.2. Критерий оценки выхода нейронов за границы области.

В данной работе таким критерием выбрана функция

$$D = \frac{\sum_E |X - P_X|^2}{N}, \tag{6}$$

где  $E$  – множество точек соответствующих нейронам сети лежащих за границей адаптируемой области,  $N$  – мощность множества  $E$ ,  $X \in E$ ,  $P_X$  - проекция  $X$  на адаптируемую область.

Функция  $D$  была построена по принципу дисперсии случайной величины,  $P_X$  – играет роль математического ожидания. Если в (6) заменить функцию проекции  $P_X$  на математическое ожидание, то получится формула величины дисперсии случайной величины, принадлежащей множеству  $E$ . Также данную функцию можно понимать как математическое ожидание квадрата расстояния от адаптируемой области до вышедшей за ее пределы точки.

Реализация модифицированного метода адаптирует области заданные пиксельной картой. Для расчета представленной оценки, граница области представляется в виде многоугольника полученного следующим образом.

1. Выделяются граничные пиксели области. Граничным пикселем считается пиксель области смежный по стороне с пикселем, не принадлежащим области (рис. 5,а).



2. Центр каждого граничного пикселя соединяется с центром смежного по стороне и углу граничного пикселя. В результате получаются один или несколько многоугольников без самопересечений (рис. 5,б).

3. Для каждого нейрона вышедшего за адаптируемую область вычисляется минимальное расстояние от всех многоугольников (рис. 5,в).

4. Квадраты полученных расстояний складываются и делятся на количество точек вышедших за границу адаптируемой области.

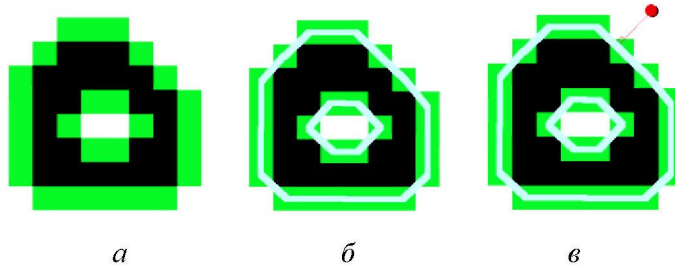


Рис. 5. Представление граничных пикселей области  
а – выделение граничных пикселей области, б – выделение многоугольников, в – расчет минимального расстояния

### 3.3. Улучшенный композиционный алгоритм.

1. Инициализация положений узлов сетки. Вычисление сеточного расстояния.
2. На первой макроитерации ( $s=1$ ) применяется базовый алгоритм в течение  $n_0$  итераций ко всем узлам сетки.
3. Удаление узлов вышедших за границу области.
4. Пересчет сеточного расстояния.
5. На каждой макроитерации с номером  $s > 1$  выполняются следующие действия:
  - а) Применение базового алгоритма в течение  $n_1(s)$  итераций к граничным узлам сетки с генерацией точки только на границе области.
  - б) Применение базового алгоритма в течение  $n_2(s)$  итераций ко всем узлам с генерацией точки во всей области. При этом все граничные узлы зафиксированы и не меняют своего положения. Кроме того, если узлом победителем является граничный узел сетки, то он заменяет случайную точку  $x(n)$ .
6. Повторяются макроитерации до тех пор, пока изменения положений узлов не станут достаточно малыми.

## 4. Программная реализация и апробация

Для разработки системы был выбран язык программирования C++. Интерфейс приложения и графика реализованы с использованием фреймворк Qt4, а также библиотеки Boost.

Пользовательский интерфейс состоит из виджета отображения сети и адаптируемой области слева. Справа расположена панель управления программой (рис. 6).

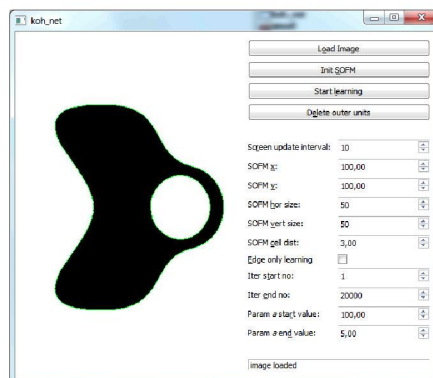


Рис. 6. Интерфейс программы

После загрузки изображения в программу производится его начальная обработка

1. выделение внутренних точек области;
2. выделение граничных точек области.

Точкой принадлежащей области считается любая точка не белого цвета, т.е. в модели RGB не равная 0хFFFFFF.

Граничным пикселем считается пиксель области смежный по стороне с пикселем, не принадлежащим области. Все остальные пиксели считаются внутренними.

Выделение точек заключается в их окраске в цветовой код для наглядности, и занесение в отдельные массивы для обеспечения работы алгоритма по генерации случайной точки.

Адаптируемая область представляет собой пиксельную карту. После загрузки изображения в программу все пиксели не белого цвета считаются принадлежащими области, и заносятся в линейный массив.

Для получения случайной точки генерируется случайный индекс массива, по этому индексу из массива берётся точка области.

Для генерации случайной величины была использована функция равномерного распределения из библиотеки *boost::random*.

Некоторые результаты построения адаптивных сеток на сложных нерегулярных областях представлены на рисунке 7.

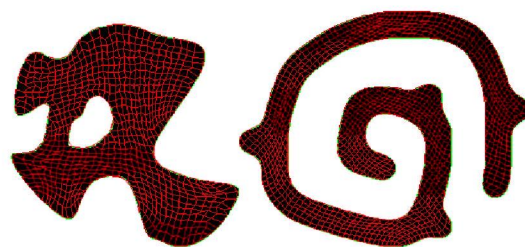


Рис. 7. Результат применения разработанного метода для построения адаптивной сетки на нерегулярной области

### Заключение

Существующие нейросетевые алгоритмы доказывают возможность построения адаптивных сеток на сложных физических областях. Результат адаптации сетки зависит от топологии и параметров обучения нейросети.

В работе апробирован подход, заключающийся в изменении топологии сети в процессе адаптации и соответствующем изменении сеточного расстояния. Разработан модифицированный алгоритм. Для этого была решена задача пересчета сеточного расстояния в новой топологии. Реализована программа для построения адаптивных сеток на нерегулярных областях, позволяющая интерактивно изменять параметры обучения сети, упрощая тем самым проведение численных экспериментов. По результатам экспериментов модифицированный подход показал лучшую адаптацию.

Тем не менее разработанный метод не является окончательным, а напротив открывает обширные возможности для дальнейшего улучшения. Замечательным является тот факт, что выбранный способ изменения сеточного расстояния слабо зависит от способа изменения топологии сети. Это делает возможным дальнейшую модификацию по двум независимым направлениям способу изменения топологии и способу расчета сеточного расстояния.

### Литература

1. Хакимзянов Г.С., Шокин Ю.И., Барахнин В.Б., Шокина Н.Ю. Численное моделирование течений жидкости с поверхностными волнами // Новосибирск: Изд-во СО РАН, 2001. – 394 с.
2. Thompson J.F., Warsi Z.U.A., Mastin C.W. Numerical grid generation, foundations and applications // Amsterdam: North-Holland, 1985.
3. Лисейкин В.Д., Лебедев А.С., Китаева И.А. Универсальный эллиптический метод построения разностных сеток. Новосибирск: НГУ, 2004. – 266 с.
4. Gordon W.J., Thiel L.C. Transfinite mappings and their applications to grid generation // Numerical Grid Generation, Appl. Mathematics and Computation. 1982. Vol.2/3. P. 171-192.
5. Годунов С.К., Проконов Г.П. О расчетах конформных отображений и построении разностных сеток // Журн. вычисл. математики и мат. физики. 1967. Т. 7. С. 1031-1059

6. Ritter H., Martinetz T., Schulten K. Neural Computation and Self-Organizing Maps: An Introduction. New York: Addison-Wesley, 1992.
7. Manevitz L., Yousef M. Finite-Element Mesh Generation Using Self-Organizing Neural Networks. //Microcomputers in Civil Engineering 12. 1997. P. 233-250.
8. Нечаева О.И. Нейросетевой подход для построения адаптивных сеток // Нейроинформатика. 2006. Ч.2. С. 172-179.
9. Нечаева О.И. Композиционный алгоритм для построения адаптивных сеток произвольной структуры // Нейроинформатика–2007: сб. науч. тр. всерос. науч.-техн. конф. М.: МИФИ, 2007. С. 72-79.
10. Koutnik J., Mazl R., Kulich M. Building of 3d environment models for mobile robotics using self-organization // In Proc. of The 9th International Conference on Parallel Problem Solving From Nature – PPSN-IX, Springer. 2006. P. 721-730.
11. Kohonen T. Self-organizing Maps // Springer Series in Information Sciences, V.30, Springer, Berlin, Heidelberg, New York, 2001. 501 p.
12. Kohonen T. K. Self-organization and associative memory. New York: Springer Verlag, 1989. – 312 p.
13. Arkhipov S.V. Neural net algorithms of construction of adaptive grid // Mongolian Mathematical Journal. 2011. Vol. 15. P. 20-29.
14. Архипов С.В. Модифицированные алгоритмы построения нейронной сети SOFM // Вестник БГУ. Вып. 9. Математика и информатика. 2011. С. 61-68.

*Архипов Сергей Валерьевич*, канд. техн. наук, доц., директор центра информационных технологий Бурятского государственного университета, тел. (3012) 297160 доб.260, e-mail: svarkh@mail.ru

*Брагин Александр Федорович*, магистр Института математики и информатики Бурятского государственного университета

*Arkhipov Sergey Valeryevich*, candidate of technical science, docent, director of the Center of information systems of the Buryat state university.

*Bragin Alexander Fedorovich*, master of Institute of mathematics and informatics of the Buryat state university

УДК 517.518.823

©Б. Болдсайхан, Т.Г. Дармаев, Б.В. Хабитуев, Ф.В. Хандаров

**МОДЕЛИРОВАНИЕ ПУЛЬСОВОЙ ВОЛНЫ ПО ХАОТИЧЕСКИМ РЯДАМ<sup>1</sup>**

В работе рассматривается подход к моделированию пульсовой волны с использованием реконструкции обыкновенных дифференциальных уравнений по хаотическим рядам.

**Ключевые слова:** традиционная медицина; математическое моделирование; динамические системы.

*B. Boldsaikhan, T.G. Darmaev, B.V. Khabituev, F.V. Khandarov*

**MODELING OF PULSE WAVE USING THE CHAOTIC SERIES**

The paper deals with an approach to modeling of the pulse wave using ODE reconstruction for the chaotic series.

**Key words:** traditional medicine, mathematical modeling, dynamical systems.

**1. Введение**

В последнее время наблюдается тенденция развития медицины превентивного действия. Главной чертой методов западной медицины является то, что процесс постановки диагноза в ней происходит последовательно, а не параллельно, и поэтому требует длительного временного отрезка для обследования пациента. В связи с этим практический интерес представляет пульсовая диагностика, в частности, методы восточной медицины. Интерес к исследованию пульсовой волны вызван тем, что в ней закодирована информация о процессах, протекающих не только в сердце и кровеносной системе, но и в различных функциональных системах организма.

Большинство известных математических моделей кровотока основаны на «резервуарной» модели Отто Франка опубликованной в конце 19 века [1]. Развитие данной идеи привело к появлению ряда интересных работ [2,3], однако недостатком данных моделей является то, что эти модели являются довольно «общими» и сложно применимыми на практике.

Другой подход при моделировании кровотока заключается в использовании теории гидродинамики [4]. При использовании этого подхода возникает множество проблем: кровь сама по себе является не Ньютонской жидкостью, течение крови происходит по сосудам и венам, нужно учитывать регуляционные функции организма и т.д. Отчасти ввиду этих причин большинство подобных моделей довольно сложны, и кроме того зачастую они применимы лишь при очень серьезных ограничениях (например модель, описанная в работе [4] применима только для небольших участков вены).

В данной работе рассматривается способ построения динамической модели пульсовой волны по хаотическим рядам.

**2. Построение динамической модели пульсовой волны по хаотическим рядам**

Коллективом авторов была создана и внедрена информационная система диагностики функционального состояния организма человека [5]. Одной из заявленных возможностей системы является работа с различными медицинскими диагностическими аппаратами, в частности с аппаратами позволяющими регистрировать сигнал пульсовой волны [6]. В ходе внедрения системы авторами была создана и наполнена база показаний данных приборов.

Сигнал пульсовой волны представляет собой замеры показаний датчиков в нескольких точках. Таким образом, показания прибора можно рассматривать как временной векторный ряд, где каждый замер является вектором состояния объекта. В работах [7,8] рассматриваются подходы и основные проблемы построения динамических моделей по временным рядам.

Искомую модель представим в виде:

$$\frac{dx}{dt} = F(x, t)$$

где  $x$  – вектор состояния объекта.

Для определения функции  $F$  используем первые  $N$  значений вектора состояния. В таком случае

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проект 11-07-92202-Монг\_a)

процедура построения модели необходимо решить следующие задачи:

- по временному ряду  $x(t_i)$  рассчитывается временной ряд  $\left\{ \frac{dx(t_i)}{dt} \right\}$ ;
- подбирается функция  $F(x, t)$ , аппроксимирующая зависимости  $\left\{ \frac{dx(t_i)}{dt} \right\}$  от  $x$ .

Первая задача решается путём численного дифференцирования, причём при наличии шумов необходимо использовать специальные методы.

Стандартным решением второй задачи является представление функции  $F$  в виде полинома  $K$ -той степени. Коэффициенты полинома вычисляются при помощи метода наименьших квадратов.

### 3. Модель пульсовой волны

При построении модели допущены следующие упрощения:

- рассматриваются только первые 3 компоненты вектора  $x$  (показания датчиков с левой руки);

- функция  $F$  ищется в виде:  $F_j = \prod_{i=1}^3 (a_i^j + b_i^j x_i)$ ;

- для определения коэффициентов функции  $F$  рассматриваются первые 201 измерение.

Для расчёта коэффициентов искомой функции использован алгоритм основанный на нелинейной регрессии по методу наименьших квадратов с использованием метода Гаусса – Ньютона [9].

Исходный ряд представляет собой показания датчиков в точках Цон, Кан, Чаг за один сердечный цикл (рис. 1) [10]. На рисунке отчётливо видна фаза систолы ( $t$  от 0 до 60) и фаза диастолы ( $t$  от 60 до 200).

Наиболее удачные результаты были получены при следующем начальном приближении искомых коэффициентов: [1 1 0.5 0.5 0.5 0.5; 0.5 0.5 1 1 0.5 0.5; 0.5 0.5 0.5 0.5 1 1].

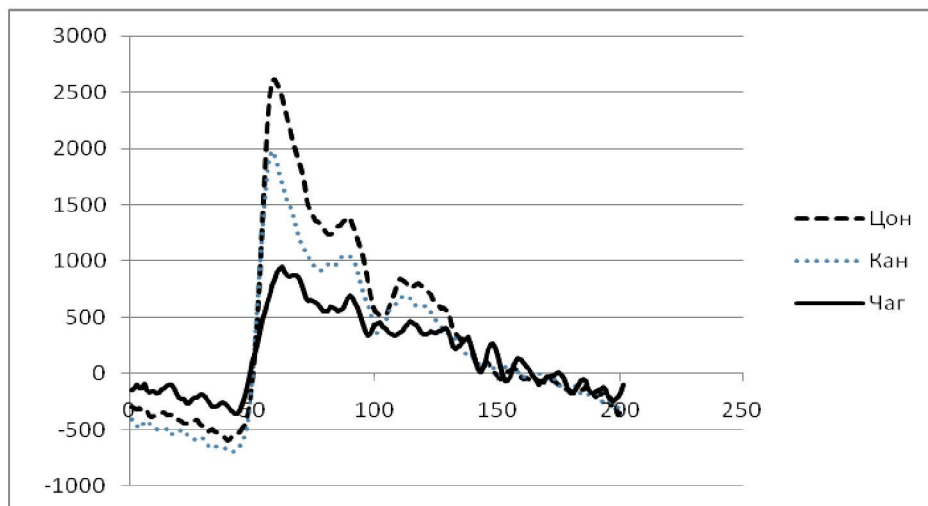


Рис. 1. Показания датчиков в точках Цон, Кан, Чаг

Значения коэффициентов полученной модели представлены в таблице 1.

Таблица 1

Значения коэффициентов модели

j	a [1,j]	b[1,j]	a [2,j]	b[2,j]	a [3,j]	b[3,j]
1	1.0000	-1.0049	1.0156	0.0028	-3.2877	0.0043
2	0.5000	-0.9969	-0.3438	-0.0009	-11.8026	0.0165
3	0.5000	-0.9934	-0.8158	0.0006	1.1124	0.0029



Полученную систему дифференциальных уравнений достаточно сложно решить аналитически, кроме того, для оценки решения нам нужны лишь значения в расчётных точках. Система была проинтегрирована при помощи метода Рунге – Кутты четвёртого порядка с начальными условиями:

$$x(0) = -300, y(0) = -407, z(0) = -150$$

Здесь  $x$  – первая координата вектора,  $y$  – вторая координата вектора,  $z$  – третья координата вектора.

Результаты численного интегрирования представлены ниже (рис. 2, рис. 3)

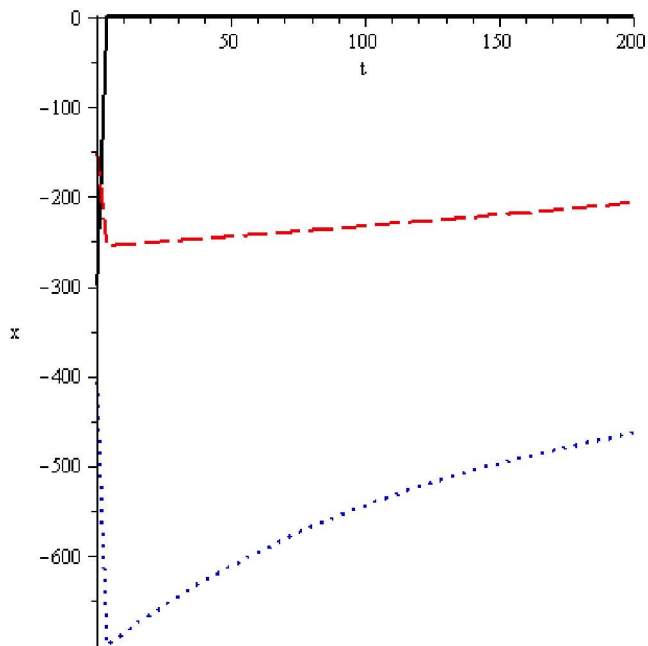


Рис. 2. Численное решение системы  
Сплошная линия –  $x(t)$ , штрих –  $y(t)$ , точки –  $z(t)$

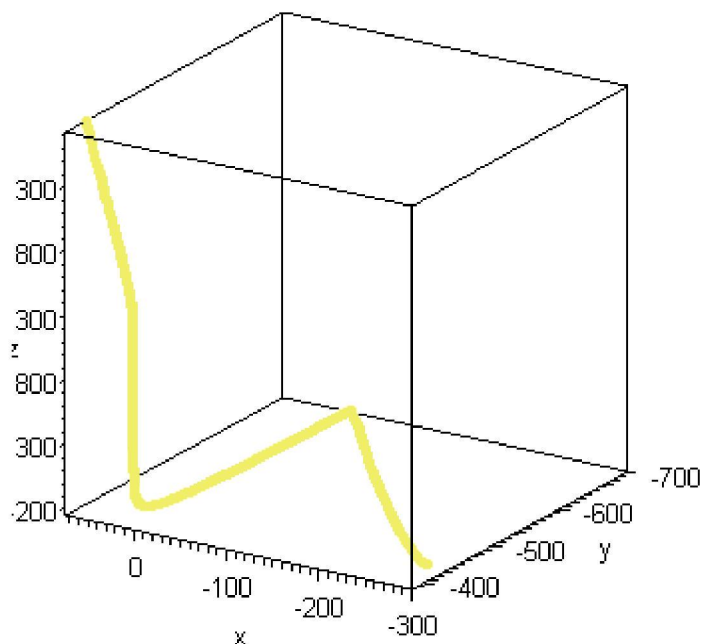


Рис.3. Фазовый портрет системы

Большинство моделей полученных в ходе численных экспериментов достаточно хорошо моделируют фазу систолы. Однако, стоит отметить, что в целом полученные модели не показывают качественных улучшений по сравнению с другими моделями [1,3,4]. Кроме того, плохо моделируется фаза диастолы и в целом модель недостаточно хорошо аппроксимирует исходные данные. Эти проблемы возникли по ряду причин:

- в качестве исходного ряда был использован реальный сигнал, который представляет собой замеры пульса за один сердечный цикл;

- вид функции  $F(x,t)$ ;

- способ нахождения производной исходного ряда.

Таким образом, необходимо решить следующие задачи:

- фильтрация исходного сигнала, исходный сигнал является реальным, немодельным сигналом, кроме того, нас интересуют достаточно продолжительные реализации, содержащие от 200 сердечных циклов. Таким образом, необходима фильтрация начального сигнала;

- нахождение вида функции  $F(x,t)$ . Если моделирование фазы систолы не составляет особого труда (большинством исследователей рассматривают её как экспоненту), то моделирование фазы диастолы – достаточно сложная задача. Вполне вероятно, что функцию  $F(x,t)$  нужно брать в достаточно сложном виде, задачу подбора вида функции можно решить в частности используя аппарат нейронных сетей при наличии достаточно продолжительных реализаций исходного сигнала;

- нахождение производной реального сигнала. Для поиска производной подобного сигнала необходимо использовать нестандартные методы аппроксимации [8].

### Заключение

Моделирование процесса кровотока и пульсовой волны в частности – сложная задача. Большинство существующих моделей математически сложны и неудобны для применения на практике. В тоже время исследование подобных моделей важны для практики.

В работе рассматривается подход, к моделированию пульсовой волны основанный на методе реконструкции обыкновенных дифференциальных уравнений по хаотическим рядам. Приведены результаты тестирования подхода на реальных данных, рассмотрены основные проблемы и способы их решения.

### Литература

1. Рашевски Н. Некоторые аспекты математической биологии. М.: Медицина, 1966. – С. 236.
2. Хабитуев Б.В., Цыбиков А.С. Модификация модели кровеносной системы Франка // Математика, ее приложения и математическое образование: III Всероссийская конф. с междунар. участием. – Улан-Удэ, 2008. – С. 365–373.
3. Кантор В.Я., Кунделев А.Ю. Применение моделей упругого резервуара при расчете параметров кровеносной системы // Проблемы машиностроения. – Харьков, 2002. – Т. 3. – С. 71-75.
4. Волобуев А.Н. Течение жидкости в трубках с эластичными стенками // Успехи физических наук. Российская академия наук. – 1995. – Т. 165, №2. – С. 177-187.
5. Информационная система функциональной диагностики с использованием методов тибетской и монгольской медицины / Т.Г. Дармаев и др. // Инфокоммуникационные образовательные технологии: модели, методы, средства, ресурсы: материалы II Байкальской межрегион. науч.-практ. конф. с междунар. участием (г. Улан-Удэ – с. Максимиха, оз. Байкал, 23-25 июня 2011 г.). – Улан-Удэ, 2011. – С.38-45.
6. Дудин С.А. Система диагностики и коррекции организма человека // Методы и алгоритмы принятия эффективных решений: сб. тр. Междунар. науч. конф. (ТТИ ЮФУ (ТРТУ)). Таганрог, 2009.
7. Безручко Б.П., Смирнов Д.А. Современные проблемы моделирования по временным рядам // Известия Саратовского университета. Сер. Физика. – Саратов, 2006. – С. 3–27.
8. Безручко Б.П. Моделирование по временным рядам в приложении к обработке сложных сигналов // Известия высших учебных заведений. Обзоры актуальных проблем нелинейной динамики. – Саратов, 2009. – Т. 17, №4. – С. 193–207.
9. Нелинейная регрессия по методу наименьших квадратов с использованием метода Гаусса-Ньютона: [Электронный ресурс]. URL: <http://matlab.exponenta.ru/statist/book2/12/nlinfit.php>
10. Чжуд-ши. Канон тибетской медицины / пер. с тиб. яз. Д.Б. Дашиева. – М.: Восточная литература РАН, 2001. – 768 с.

*Болдсайхан Бадамжав*, д-р мед. наук, проф.; Монгольский университет науки и технологий; директор Центра системных исследований Монгольского университета науки и технологий; 210646, Улан-Батор, ул. Бага Тойруу, 34; 99133579; e-mail: Bbld51@yahoo.com

*Дармаев Тумэн Гомбоцыренович*; канд. физ.-мат. наук; доц.; директор Научно-образовательного и инновационного центра системных исследований и автоматизации; 670000, г. Улан-Удэ, ул. Смолина, 24 «а»; (3012)221215; e-mail: dtg@bsu.ru

*Хабитиев Баир Викторович*; инженер Научно-образовательного и инновационного центра системных исследований и автоматизации; 670000, г. Улан-Удэ, ул. Смолина, 24 «а»; (3012)221215; e-mail: bairinc0@gmail.com

*Хандаров Фёдор Владимирович* – научный сотрудник Научно-образовательного и инновационного центра системных исследований и автоматизации; 670000, г. Улан-Удэ, ул. Смолина, 24 «а»; (3012)221215; e-mail: fhandar@rambler.ru

*Boldsaikhan Badamjav*; Ph. D., Sc.D., M.D., prof.; Director of system science research center, Mongolian university of science and technology; 210646, Ulaanbaatar, Baga Toiruu 34; 99133579; e-mail: Bbld51@yahoo.com

*Darmaev Tumen Gombotsyrenovitch*; Ph.D. in math; Associate Professor; Director of scientific and educational centre of system research and automatization; 670031, Ulan-Ude, ul. Smolina , 24 «а» ; (3012)221215; e-mail: dtg@bsu.ru

*Khabituev Bair Victorovitch* – Software Engineer of scientific and educational centre of system research and automatization; 670042, Ulan-Ude, Ulan-Ude, ul. Smolina , 24 «а»; (3012)221215; e-mail: bairinc0@gmail.com

*Khandarov Fedor Vladimirovitch* – Research associate of scientific and educational centre of system research and automatization; 670000, Ulan-Ude, ul. Smolina , 24 «а»; (3012)221215; e-mail: fhandar@rambler.ru



УДК 519.83+115.330

©А.В. Воротынцев

## НЕКОТОРЫЕ ХАРАКТЕРНЫЕ ГРАФЫ РАБОТ ДЛЯ ПРЕДСТАВЛЕНИЯ СЦЕНАРИЕВ РАСЧЕТА МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

Обсуждаются представления графами работ сценариев расчета характерных вычислимых на компьютере математических моделей. Представления ориентированы на расчеты, выполняемые в сети Internet с помощью разработанной сетевой библиотеки вычислимых моделей.

**Ключевые слова:** распределенные вычисления, «облачные» вычисления, графы работ, представление моделей, математическое моделирование.

*A.V. Vorotyntsev*

## SOME TYPICAL WORKFLOW GRAPHS FOR REPRESENTATION OF THE SCENARIOS OF CALCULATING THE MATHEMATICAL MODELS

The representations by workflow graphs of calculation scenarios of typical computable models are discussed. The representations are focused on the calculations which are carried out in Internet network by means of developed network library of computable models.

**Keywords:** distributed calculations, cloud computing, workflow graphs, representation of models, network libraries, mathematical modeling.

### Введение

В связи с нарастающим многообразием и сложностью знаний становится актуальным создание сетевых библиотек моделей. В [1–4] изложены концепция и архитектура сетевой библиотеки Нива вычислимых математических моделей, результаты которых получают вычислениями на удаленном компьютере. К таким моделям относятся, например, имитационные модели.

Задача сетевой библиотеки – сохранять модели и предоставлять легкодоступные услуги использования моделей, прежде всего сложных, широкому кругу сетевых пользователей, как правило, не обладающих полным знанием модели. К таковым обычно относятся специалисты предметной области, для которых создаются модели.

В сетевой библиотеке модели представляются сценариями их расчета и, возможно, иного использования. Сценарий для расчета модели на удаленном компьютере пользователь самостоятельно конструирует в виде графа работ модельных компонент на своем компьютере. Это понуждает применять наиболее простые и безопасные представления моделей. Поэтому концепция сетевой библиотеки ориентирована на графическое представление пользователю структуры модели и раскрытие ее второстепенных деталей. Далее удаленный компьютер называется сервером. Сервер рассчитывает сценарий и результаты пересылает пользователю.

Графы работ выбраны для представления моделей из-за их наглядности и универсальности. Они позволяют наглядно представлять многообразие сценариев расчета моделей с помощью композиций модельных компонент и организовать их хранение в базе данных сетевой библиотеки на сервере. Задавая различные входные данные, а также конструируя разные графы работ модельных компонент, пользователь может варьировать расчетную модель в предусмотренных пределах. Отметим, что модельной компонентой может быть исполняемый код любой степени сложности и назначения, причем даже не связанный с математическими вычислениями, например, интерпретатор или вычислитель символьных выражений.

Библиотека не передает пользователю код модели. С согласия автора библиотека может предоставлять пользователю текстовое описание модели. Автор сам решает, какие сведения о модели можно поместить в ее описание. Это позволяет сохранять авторские права.

Всем этим сетевая библиотека существенно отличается от известных программ MathLab, Maple и т.д., ориентированных более на исследователей своих собственных моделей, а не их пользователей.

Представление задач графами работ хорошо известно и используется в программировании, например, [5]. Однако систематические применения таких представлений к задачам математического моделирования автору неизвестны. Поэтому представляет интерес исследование вида графов работ, которые могут быть использованы для представления характерных вычислительных задач

моделирования. В статье обсуждаются графы работ для представления деревьев выражений, деревьев рекурсии и задач, решаемых итерациями, на примере задачи Коши, а также визуальное конструирование систем уравнений. За исключением графа работ для задачи Коши, эти вопросы ранее практически не обсуждались.

### 1. Графы работ и компоненты

Компонентой  $M=\{F,D\}$  вычислимой модели называется исполняемый компьютером программный код, который реализует некоторую функцию  $F$  без параметров и исполняет ее над набором  $D$  специфических данных, называемых далее слотами. Компоненты, имеющие одни и те же  $F$  и  $D$ , но отличающиеся значениями данных из  $D$ , назовем МодОбъектами. Композицию МодОбъектов, представленных узлами в графе работ, назовем сценарием использования модели. Совокупность сценариев использования одной модели будем называть ГиперМоделью.

Каждому МодОбъекту приписывается содержательное имя вида  $[m][d]$ , где  $m$  – имя для обозначения функциональности  $F$  компоненты  $M=\{F,D\}$ , а  $d$  – имя набора значений слотов  $D$  этой компоненты. МодОбъекты вместе с данными их слотов сохраняются в базе данных сервера. Имена  $[m][d]$  используются для поиска МодОбъектов в базе данных и в графе работ. На графе можно не указывать  $[d]$ , если компонента имеет только один МодОбъект.

Сценарий визуально (мышкой) конструируется пользователем в виде графа работ. Сценарий – это то, что и рассчитывается удаленным сервером после ввода пользователем своих данных в компоненты и проверки правильности графа сервером. Результаты расчета сервер посылает пользователю на его компьютер.

Таким образом, модели в сетевой библиотеке представляются семействами допустимых сценариев - графов работ модельных компонент. Каждое семейство допустимых графов имеет формальное описание, хранимое в библиотеке. В описаниях допустимые компоненты указываются своими именами вида  $[m][d]$ .

Под работой понимается исполнение компонентой  $\{F,D\}$  своей функции  $F$ . Работы выполняются в последовательности, указанной стрелками графа работ. В дальнейшем предполагается включение задач, решаемых параллельно.

Декомпозиция модели в граф работ поддерживается механизмом обмена данными между модельными компонентами с помощью слотов. Слотом здесь называется специальная переменная в программе компоненты, имеющая следующие свойства. Как переменная слот имеет тип и значение, например, числовое. Слот может представлять массив чисел, строку символов и т.д. Перед началом вычислений слоты инициализируются, т.е. получают начальные значения. Значения либо указываются в коде компоненты, либо задаются пользователем. Функция  $F$  компоненты  $M=\{F,D\}$  может использовать значение своего слота  $M.X$  как обычное данное, а также изменять его значение. Функция  $F$  не может непосредственно изменять слоты других компонент.

Суть в том, что слоты различных компонент могут обмениваться значениями, если они связаны. Этими обменами поддерживается равенство значений связанных слотов. Так, если слот  $M1.X$  компоненты  $M1$  связан со слотом  $M2.Y$  (это обозначается как  $M1.X = @M2.Y$ ), тогда изменение одной компонентой значения своего слота немедленно вызывает такое же изменение значения слота другой компоненты.

Таким образом, выполняя вычислительную работу, компонента может использовать результаты вычислительной работы других компонент, переданные ей через связанные слоты.

Описания допустимых конфигураций графов работ и связей между слотами компонент помещаются в библиотеку вместе с компонентами модели. Используя такие описания, исполняющая подсистема библиотеки устанавливает связи между слотами перед началом вычисления заданного сценария. Если при этом связанная с компонентой  $M1$  компонента  $M2$  не будет найдена, значение слота  $M1.X$  останется равным первоначальному значению, полученному ранее при инициализации.

Под исполняющей подсистемой сетевой библиотеки понимается та часть всей системы, которая исполняет команды пользователя, например, выполняет проверку и расчет графа работ, пересылку результатов, исполнение запросов к базе данных и т.д. Другую часть системы представляют данные, компоненты и их описания, хранимые в базе данных на сервере.

Связи слотов могут изображаться на графе работ только для иллюстрации. В статье не рассматриваются алгоритмы установления этих связей. В большинстве приводимых примеров они более или менее интуитивно ясны.

## 2. Представление характерных задач графами работ

### Представление дерева выражения

Представление арифметических выражений деревьями выражений широко используется в программировании. Такие деревья позволяют изобразить процесс вычисления выражения.

Представление произвольного дерева выражения графом работ поясним на примере простейшего выражения  $F(x)=f(x)*h(z)+g(y)$ , дерево которого показано на рис. 1.1. Узлы дерева помечены операциями. Согласно известным правилам вычисления дерева  $F$  сначала вычисляются значения узлов  $f$  и  $h$ , затем умножением получают значение узла  $*$ . Затем вычисляется значение  $g$ . Сложением значений узлов  $*$  и  $g$  получают значение выражения  $F$  в узле  $+$ .

Эти вычисления можно реализовать с помощью наборов ячеек памяти, изображенных прямоугольниками на рис. 1.2. Пусть левая ячейка каждого набора содержит результат операции, указанной в следующей ячейке набора. В нашем примере – это операции  $f, g, *, +$ . Пусть последние ячейки содержат слоты аргументов операции.

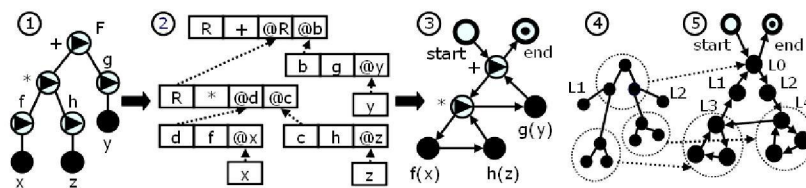


Рис. 1. Преобразование дерева выражения в граф работ

Вычислив  $d$  и  $c$  как результат функций  $f$  и  $h$  с аргументами-слотами  $@x$  и  $@z$ , получаем значения слотов  $@d$  и  $@c$ . Здесь значения  $x, z, y$  для слотов вводятся пользователем. Выполняя операцию  $*$  над аргументами-слотами  $@d$  и  $@c$ , получим результат в левой ячейке  $R$ . Согласно определению слота, приведенному выше, значение результата немедленно передается связанному слоту  $@R$ , помещенному в 3-ю ячейку операции  $+$ . Затем в ячейке  $b$  вычисляется  $g(y)$ . Результат немедленно передается связанному слоту  $@b$ . Операция  $+$  вычисляет сумму значений своих слотов  $@R$  и  $@b$ . Результат она помещает в верхнюю ячейку  $R$ .

Изобразим на рис. 1.3 каждый набор ячеек рис. 1.2 одним узлом графа, а последовательность описанных выше операций над ячейками изобразим стрелками. Будем считать работой узла выполнение операции, указанной во второй ячейке соответствующего набора. Добавим узлы  $start$  и  $end$  для обозначения начала и конца работ. Примем соглашение, что, попадая по стрелке в узел, например в узел  $*$ , следует делать сразу переход по стрелке вниз, если такая стрелка имеется, а не выполнять работу этого узла. Работа этого узла будет выполнена при возвращении в узел. Нетрудно видеть, что полученный на рис. 1.3 граф работ однозначно определяет вычисление выражения  $F$ .

Описанный подход указывает, как создать в сетевой библиотеке множество компонент, представляющих известные унарные, бинарные операции и другие элементарные функции. Из таких компонент можно конструировать графы работ для вычисления различных математических выражений. Назовем такие компоненты элементарными.

Такое конструирование можно сделать удобным для обычного пользователя, если на экране его компьютера показывать изображения компонент и позволить пользователю перетаскивать их мышкой, а затем соединять стрелками, образуя направленный граф работ. Графы работ элементарных компонент удобны также тем, что допускают простое стандартное описание.

Однако дерево выражения трудно конструировать визуально, если оно объемное. Такие деревья можно представлять более компактными деревьями кольцевых подграфов, получаемыми агрегированием исходного дерева. Не будем усложнять изложение определением таких деревьев. Представьте обычное дерево корнем вверх и замените в нем каждый узел некоторым кольцом узлов. Получится примерно такой же граф кольцевых подграфов, как показанный на рис. 1.5.

Поясним, как при агрегировании образуются кольцевые подграфы. На рис. 1 показана процедура агрегирования дерева выражения на рис. 1.4 в дерево кольцевых подграфов на рис. 1.5. Здесь три верхних элементарных узла в пунктирном эллипсе агрегируются в один узел  $L0$ . Видно, что дерево на рис. 1.5 включает три кольцевых подграфа. Поскольку в узлы верхнего эллипса переда-

вали данные четыре элементарных узла дерева, то узел L0 будет оперировать с данными, полученными от узлов L1, L2, L3, L4. Таким образом, агрегирование узлов стандартной функциональности, выполняющих элементарные операции '+', '-', '\*' и т.д., создает неэлементарный узел произвольной функциональности. Видно, что получившийся на рис. 1.5 граф работ является деревом кольцевых подграфов.

Одно и то же дерево можно агрегировать по-разному. Количество возможных неэлементарных компонент, получаемых подобным агрегированием деревьев для различных моделей, огромно. Поэтому возникает проблема компактного графического представления характерных алгоритмов и моделей деревьями кольцевых подграфов работ.

В заключение можно сделать замечания об организации связей между слотами, почерпнутыми из анализа графов работ элементарных компонент. Как правило, связи желательно делать короткими, т.е. преимущественно связывать слоты компонент внутри одного кольцевого подграфа. Данные по слотам должны перемещаться преимущественно внутри подграфа и вверх к корню дерева. Двигаясь вверх, данные обычно агрегируются.

### Представление дерева рекурсии

Рекурсии играют большую роль в теории алгоритмов. Совсем не считая исчерпанной задачу представления рекурсий графами работ, обсудим простой пример. Рассмотрим представление графом работ рекуррентной последовательности  $Y_n = 0.5 * (Y_{n-1} + A/Y_{n-1})$ ,  $Y_1 = 1.5$ ,  $n \geq 1$ , сходящейся к  $\sqrt{A}$ . Выражение для  $Y_n$  является арифметическим выражением и для него можно привести очевидное дерево выражения, аналогичное дереву на рис. 1.1. Такое дерево называется деревом рекурсии. Описанной в предыдущем разделе процедурой это дерево можно преобразовать в графы работ, показанные на рис. 2. Ромб изображает условную операцию if. Граф 2.2 отличается тем, что дерево выражения  $Y_{n-1} + A/Y_{n-1}$  представлено кольцевым подграфом, чтобы исключить повторное вычисление компоненты [Recur][n-1]. Здесь пунктирные стрелки изображают передачу связанными слотами компонент результатов вычислений.

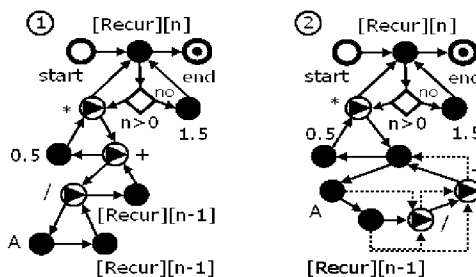


Рис. 2. Граф работ, представляющий рекуррентную последовательность  $Y_n$

Рассмотрим рекурсивную функцию Func:

```
double Func(double A, int n)
{ return if n=1 then 1.5 else 0.5*(Func(A,n-1) + A/Func(A,n-1)); }
```

вычисляющую  $Y_n$  для действительного числа  $A$  и целого  $n$ . Напомним, как вычисляется Func. Чтобы вычислить Func(A,n), программа должна сначала вычислить Func(A,n-1). Если она не может этого сделать, например при  $n > 2$ , она создает в памяти копию Func с параметрами (A,n-1), из которой вызывается Func(A,n-2). Если, вызвав Func(A,n-2), программа не может ее вычислить, создается ее копия, из которой вызывается Func(A,n-3). И так продолжается до тех пор, пока значение аргумента  $n$  в копии Func не окажется равным 1, и программа сможет вычислить Func(A,1). Теперь программа возвращается обратно к предыдущей копии и вычисляет предыдущую копию Func(A,2). Последовательно возвращаясь к начальной копии Func(A,n), она вычисляет результат. Процесс вычислений описан не совсем точно, но он достаточно поясняет работу, выполняемую компонентой Rescu в графах работ на рис. 2.

Как и функция Func, компонента Rescu создает копии подчиненного ей подграфа работ до тех пор, пока не сможет его вычислить. После этого вычисляются предыдущие подграфы в последовательности обратной к последовательности их создания копированием. Отметим, что код Rescu не зависит от подчиненного подграфа, если подграф состоит из элементарных компонент. Такую

Ресур можно также считать элементарной компонентой.

### Представление задачи Коши, решаемой методом Рунге-Кутты

Рассмотрим часто решаемую задачу Коши для системы уравнений:

$$dx/dt = F(t, x), \quad x(0) = x_0, \quad x \in R^n, \quad 0 \leq t \leq T. \quad (1)$$

Очевидно, полное графическое представление задачи Коши лишено смысла из-за его обычно большого объема. Используемая концепция позволяет представить задачу в общем простом виде циклического графа из двух узлов на рис. 3.2.

На рис. 3.1 в верхнем прямоугольнике RungeKM изображена блок-схема алгоритма Рунге-Кутты-Мерсона, а в нижнем прямоугольнике F – граф расчета правой части  $Y = F(X)$ ,  $X = (t, x)$ , уравнения (1). Компонента RungeKM графа работ на рис. 3.2 изображает блок-схему в прямоугольнике RungeKM, а компонента F – граф в прямоугольнике F.

Вычисления блок-схемы RungeKM начинаются в узле start и заканчиваются в узле end, когда время k-й итерации  $t_k > T$ . Работа алгоритма на k-ой итерации заключается в вычислении  $K_1, \dots, K_5$  – значений  $F(X)$  для 5-ти пар  $X = (t_k, X_k), \dots, (t_k + h, X_k + \dots)$ , указанных на рис. 3.1 и вычислении значения  $X_{k+1}$  решения (1) для k+1 итерации при достижении погрешностью E заданной точности  $\epsilon$  [1]. Чтобы вычислить  $K_1 = F(t_k, X_k)$ , значение  $X = (t_k, X_k)$  присваивается слоту @X и им передается в X компоненты F. Следующий узел @F блок-схемы совершает переход к компоненте F, где функция F(X) вычисляет значение слота Y=F(X), которое присваивается  $K_1$  компоненты RungeKM. Затем выполняется возврат в RungeKM и переход к вычислению  $K_2$ , повторяющему вычисление  $K_1$ . Подробности работы блок-схемы RungeKM см. в [1].

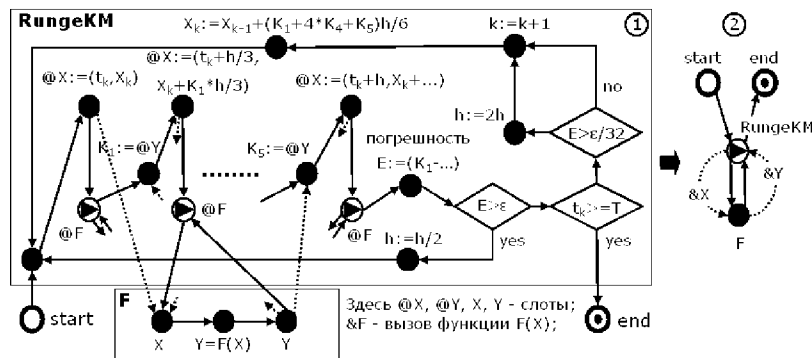


Рис. 3. Представление метода Рунге-Кутты-Мерсона графом работ

Нетрудно видеть, что граф работ на рис. 3.2 представляет множество других задач, решаемых итерациями, например, задачу нахождения корня уравнения  $F(x) = 0$  или задачу поиска экстремума функции  $F(x)$ . Для таких задач характерно наличие выделяемой компоненты-итератора, выполняющей в цикле тестирование известной функции F до тех пор, пока известный критерий, например, точность не достигнет заданного значения. Такой итератор реализуется компонентой, аналогичной RungeKM, а функция – компонентой, вычисляющей функцию F. Функцию F компоненты пользователь может задавать либо в текстовом виде, либо с помощью дерева выражения.

Правая часть системы (1) может быть реализована несколькими компонентами так, чтобы на рис. 3.2 вместо одного узла F оказалась последовательность узлов, представляющих эти компоненты. Характерным примером может служить задача массопереноса в многослойной среде, решаемая методом прямых [1]. В этом случае компонента моделирует перенос в отдельном слое. Составляя визуальную последовательность таких компонент, пользователь может рассчитывать перенос в среде с заранее неизвестным числом слоев. Таким образом, визуальное составление последовательности компонент, пользователь может решать задачи (1) различной размерности с помощью компонент RungeKM.

Дополнительные возможности визуального конструирования задачи Коши (1) пользователю дает метод уточнения, описанный в следующем разделе.

**Конструирование систем уравнений методом уточнения**

Опишем полезную процедуру, названную здесь методом уточнения. Для этого рассмотрим систему уравнений:

$$dY/dt = H, \quad H \equiv H(t, Y, @A, @B); \tag{2}$$

$$dZ/dt = G, \quad G \equiv G(t, Z, @C, @D); \tag{3}$$

относительно  $Y$  и  $Z$ , где  $@A, @B, @C, @D$  – слоты в компонентах  $H$  и  $G$ , рассчитывающих правые части уравнений (2) и (3) соответственно. Пусть  $S$  – компонента, решающая системы вида (2) и (3), например, методом Рунге-Кутты-Мерсона, как было описано выше. Пусть при инициализации перед началом вычислений слотам присваиваются некоторые начальные значения  $A_0, B_0, C_0, D_0$ , заданные пользователем.

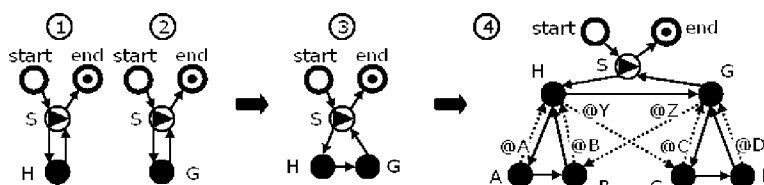


Рис. 4. Пример конструирования системы методом уточнения

Если при установлении связей между слотами, исполняющая подсистема библиотеки не обнаружит наличия в графе работ компонент  $A, B$  и  $C, D$ , подчиненных соответственно узлам  $H$  и  $G$ , значения слотов, как указывалось в разделе 1 «Графы работ и компоненты», не изменятся и останутся равными значениям, полученным при инициализации. В этом случае система (2)–(3) будет состоять из двух независимых подсистем (4)–(5) с постоянными  $@A = A_0, @B = B_0, @C = C_0, @D = D_0$ . Ее представляет граф работ на рис. 4.3.

$$dY/dt = H(t, Y, A_0, B_0), \tag{4}$$

$$dZ/dt = G(t, Z, C_0, D_0). \tag{5}$$

Пусть теперь пользователь включил в граф работ, как показано на рис. 4.4, узлы  $A, B$  и  $C, D$ , вычисляющие, например, функции  $A(Y), B(Y, @Z), C(@Y, Z), D(Z)$  соответственно. Тогда после инициализации слотов начальными значениями исполняющая система библиотеки обнаружит подчиненные  $H$  и  $G$  узлы  $A, B$  и  $C, D$  и установит связи между их слотами, как показано на рис. 4.4 пунктирными стрелками. В результате система уравнений окажется приведенной к виду (2)–(3) с параметрами  $@A = A(Y), @B = B(Y, Z), @C = C(Y, Z), @D = D(Z)$ , т.е. к сопряженной системе уравнений

$$dY/dt = H(t, Y, A(Y), B(Y, Z)), \tag{6}$$

$$dZ/dt = G(t, Z, C(Y, Z), D(Z)). \tag{7}$$

Из этой системы видно, что роли узлов  $A, D$  несколько отличаются от ролей узлов  $B, C$ . Узлы  $A$  и  $D$  не делают систему (2)–(3) сопряженной, а только изменяют функциональность каждого уравнения  $H$  и  $G$  по отдельности. Напротив, узлы  $B$  и  $C$ , устанавливая обмен значениями переменных из разных уравнений, связывают их в сопряженную систему уравнений. Такое изменение функциональности узлов подчиненными узлами будем называть уточнением исходной модели (4)–(5). Уточнение исходной модели, выполняемое узлами  $A, D$  будем называть нисходящим уточнением, а уточнение путем сопряжения разных задач – уточнением сопряжения.

Нисходящее уточнение позволяет осмысленно организовать иерархию свойств модели. В такой иерархии верхнее положение занимает компонента, реализующая самые общие свойства модели, допускающие нисходящее уточнение своих параметров и функциональности подчиненными компонентами второго уровня. Таких компонент может быть несколько, соединенных в графе последовательностью своих работ. В свою очередь компоненты второго уровня могут содержать параметры, нисходящим образом уточняемые подчиненными им компонентами третьего уровня. Заметим, что здесь просматриваются интересные аналогии с нисходящим уточнением функциональности в иерархиях объектов, широко используемым в объектно-ориентированных языках программирования для создания библиотек компонентов.

Уточнение сопряжением позволяет сопрягать разные задачи, имеющиеся в сетевой библиотеке.

Интерес к этой возможности также обусловлен аналогией со ссылками на сопутствующие тексты, которые широко используются в библиотеках текстов.

Нисходящие уточнение одной модели и сопряжение разных моделей относятся к фундаментальным свойствам обсуждаемых графов работ.

### **Заключение**

Описаны графы работ для расчета характерных задач математического моделирования, представленных деревом выражения, деревом рекурсии и задачей Коши. Предложены методы нисходящего уточнения и уточнения сопряжением для построения иерархий и комбинирования характерных задач.

Благодарю проф. В.А. Серебрякова, чл.-кор. РАН И.Г. Поспелова и Ю.Н. Павловского, проф. А.М. Тарко, канд. физ.-мат. наук А.А. Бездушного (ВЦ РАН), д-ра физ.-мат. наук Л.Н. Столярова (МФТИ) за обсуждение, критические замечания и внимание к работе.

### **Литература**

1. Воротынцев А.В. Экологические модели как сценарии в сетевой библиотеке вычислимых моделей. Пример публикации и расчетов двухфазной модели переноса почвенной влаги // Известия Самарского научного центра РАН, спецвыпуск по итогам национальной конференции с международным участием ЭКОМАТМОД – 2009. 2009. Т. 11, № 1(7). С. 1655–1660.
2. Воротынцев А.В. Концепция сетевых информационно-вычислительных библиотек моделей. – М.: Вычислительный центр РАН, 2009. – 109 с.
3. Воротынцев А.В. О представлении вычислимых моделей графами работ // Математическое моделирование развивающейся экономики, экологии и биотехнологий – ЭКОМОД–2010: тр. V Всерос. науч. конф., г. Киров, 5–11 июля 2010. Киров: Изд-во ВятГУ, 2010. С. 45–54.
4. Воротынцев А.В. Представление вычислимых моделей иерархическими автоматами в сетевой библиотеке // Вестник БГУ. 2011. Вып. 9. С. 87-88.
5. Кознов Д.В. Основы визуального моделирования. – М.: Интернет-университет информационных технологий. БИНОМ Лаборатория знаний, 2008. – 246 с.

*Воротынцев Александр Васильевич*, канд. физ.-мат. наук, научный сотрудник ВЦ РАН, 119333, Москва, ул. Вавилова, 40; тел. 8965-1353931, e-mail: avv\_alexv@mail.ru.

*Vorotyntsev Alexander Vasilievich*, candidate of physical and mathematical sciences, research fellow of CCAS, Vavilov st. 40, 117967 Moscow GSP-1, Russia; phone +7965-1353931, e-mail: avv\_alexv@mail.ru.



556.013:556.536.2:556.555

©В.Н. Данчев, П.Ю. Пушистов

**ОПЫТ РАЗРАБОТКИ И РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ  
ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА  
ДЛЯ МОДЕЛИРОВАНИЯ ГИДРОДИНАМИКИ И КАЧЕСТВА ВОДЫ РЕК  
И ВОДОЕМОВ БАСЕЙНА ОБИ. ЧАСТЬ 2 - ТЕЛЕЦКОЕ ОЗЕРО  
И УЧАСТОК РЕКИ ЧУЛЫШМАН**

Описаны цели разработки, структура и технология применения информационно-вычислительного комплекса для моделирования динамики экосистемы Телецкого озера. Обсуждаются проблемы калибровки и верификации комплекса, представлены результаты численных экспериментов по моделированию циркуляции и термического режима южной части Телецкого озера в весенний период.

**Ключевые слова:** Телецкое озеро, водная экосистема, модель гидродинамики и качества воды, база данных, речной термический бар.

*V.N. Danchev, P.Yu. Pushistov*

**EXPERIENCE AND RESULTS OF THE DEVELOPMENT OF INFORMATION  
AND COMPUTATIONAL COMPLEX FOR SIMULATION OF HYDRODYNAMICS  
AND WATER QUALITY OF RIVERS AND LAKES OB BASIN.  
PART 2 - LAKE TELETSKOYE AND SECTION OF THE RIVER CHULYSHMAN**

Objectives of the development, structure and technology of information and computer system for modeling the dynamics of the ecosystem of Lake Teletskoye are described. The problems of calibration and verification of the complex, the results of numerical experiments on modeling of circulation and thermal regime of the southern part of Lake Teletskoye in the spring are discussed.

**Keywords:** Teletskoye Lake, aquatic ecosystem, the model of hydrodynamics and water quality, data base, riverine thermal bar.

**Введение**

Водными объектами, для которых разработан, описываемый в данной статье информационно-вычислительный комплекс (ИВК) моделирования гидродинамики и качества воды, являются Телецкое озеро и устьевой участок реки Чулышман. Телецкое озеро – крупнейший водоем бассейна реки Оби [1-3]. Озеро состоит из двух частей: северной субширотной, длиной около 37 км (включая Камгинский залив) и субмеридиональной, вытянутой на 48 км. Средняя и максимальная ширина озера равна 2,9 и 5,2 км, соответственно, площадь зеркала — 227,3 км<sup>2</sup>. Благодаря большой глубине (средняя глубина озера - 181 м, максимальная - 323 м) в нём содержится более 40 км<sup>3</sup> пресной воды высокого качества, прозрачность которой достигает 12—15 метров. Впадает в озеро около 70 рек и 150 временных водотоков, причем более 70 % всей поступающей воды даёт река Чулышман. Вытекает из Телецкого озера река Бия, которая, сливаясь с Катунью, образует реку Обь. Телецкое озеро в 1998 г. внесено ЮНЕСКО в реестр объектов Всемирного природного наследия [3].

История исследования Телецкого озера берет свое начало в XVIII веке и подробно описана в [1-4]. Значительной вехой в экспериментальном изучении озера является 1965 год, когда на базе метеорологической станции поселка Яйлю была открыта Озерная станция Западно-Сибирского УГМС. По результатам комплексных наблюдений Озерной станции за период с 1965 по 1980 гг. авторам [1] удалось создать основу феноменологической модели гидрометеорологического режима Телецкого озера, которая активно используется до настоящего времени, в том числе, при построении и верификации двумерных моделей циркуляции и термического режима озера [5, 6 и др.].

В Институте водных и экологических проблем СО РАН, начиная с 1989 г., реализуется комплексная программа, включающая гидрофизические, гидрохимические, гидробиологические и палеолимнологические исследования. Цель исследований - установление направления и скорости сукцессии экосистемы Телецкого озера во внутригодовом и многолетнем аспектах, как результата взаимодействия физических, химических и биологических процессов, природных и антропогенных факторов [4].



Что касается антропогенных нагрузок на экосистему и водосборный бассейн озера, то в [7] на основе анализа использования природных систем выделено 5 классов геосистем по преобладающему типу антропогенного воздействия: селитебные, горнопромышленные, лесохозяйственные, сельскохозяйственные, рекреационные. Научно обоснованная система оперативного контроля и управления антропогенной нагрузкой на акваторию и бассейн Телецкого озера отсутствует [3].

Таким образом, проблема разработки методики расчета допустимых антропогенных нагрузок на экосистему Телецкого озера является актуальной, и решать ее, по-видимому, целесообразно в рамках проекта создания системы поддержки принятия решения при интегрированном управлении бассейном озера [8]. Пример успешного и инновационного, по своей сути, проекта обоснования норм антропогенных нагрузок (TMDL – суммарная максимальная суточная нагрузка) на экосистему глубокого горного озера Ватком приведен в [9,10]. Инструментом анализа TMDL озера Ватком является единая информационно-моделирующая система (ИМС), включающая в себя модель гидродинамики и качества воды CE-QUAL-W2 (далее W2) [9], для описания динамики экосистемы озера и модель экосистемы бассейна озера - HSFP (фортран-программа гидрологического моделирования) [10].

Близкий по замыслу проект, дополнительно предполагающий включение в состав ИМС численных моделей динамики погодно-климатического форсинга на площади водосбора Телецкого озера, представлен в [11]. Цель настоящей статьи – ознакомить читателя с первыми результатами разработки ИВК для описания динамики экосистемы Телецкого озера с прямым учетом гидродинамики и качества воды устьевого участка реки Чулышман (далее, ИВК «Телецкое озеро»).

### 1. Состав информационно-вычислительного комплекса «Телецкое озеро»

Информационно-вычислительный комплекс, разработанный для моделирования динамики экосистемы Телецкого озера и гидродинамики и качества воды устьевого участка реки Чулышман, включает в себя три модуля: 1 – система моделирования бассейна – WMS; 2 - версия 3.5 модели W2 (далее - W2 v.3.5); 3 – БД для обеспечения работы модулей 1 и 2.

Система базовых уравнений модели W2 v.3.5 имеет следующий вид [12]:

$$\frac{\partial UB}{\partial t} + \frac{\partial UUB}{\partial x} + \frac{\partial WUB}{\partial z} = gB \sin \alpha + g \cos \alpha B \frac{\partial \eta}{\partial x} - \frac{g \cos \alpha B}{\rho} \int_{\eta}^z \frac{\partial \rho}{\partial x} dz + \frac{1}{\rho} \frac{\partial B \tau_x}{\partial x} + \frac{1}{\rho} \frac{\partial B \tau_z}{\partial z} + qBU_x, \quad (1)$$

$$0 = g \cos \alpha - \frac{1}{\rho} \frac{\partial P}{\partial z}, \quad (2)$$

$$\frac{\partial UB}{\partial x} + \frac{\partial WB}{\partial z} = qB, \quad (3)$$

$$\rho = f(T_w, \Phi_{TDS}, \Phi_{ISS}), \quad (4)$$

$$B_{\eta} \frac{\partial \eta}{\partial t} = \frac{\partial}{\partial x} \int_{\eta}^h UB ds - \int_{\eta}^h qB ds, \quad (5)$$

где  $t$  – время;  $U$  и  $W$  – компоненты скорости вдоль осей  $x$  и  $z$ , соответственно;  $B$  – ширина русла реки или котловины озера;  $g$  – ускорение силы тяжести;  $P$  – давление;  $\tau_x$  – поперечно осредненное касательное напряжение вдоль оси  $x$ ;  $\tau_z$  – поперечно осредненное касательное напряжение вдоль оси  $z$ ;  $\rho$  – плотность;  $\eta$  – уровень поверхности воды;  $q$  – распределение по  $x$  и  $z$  результирующих расходов воды боковых притоков;  $f(T_w, \Phi_{TDS}, \Phi_{ISS})$  – функция зависимости плотности от температуры воды ( $T_w$ ), солености или суммарной концентрации растворенных в воде веществ ( $\Phi_{TDS}$ ), и концентрации неорганических взвешенных веществ ( $\Phi_{ISS}$ );  $\alpha$  – угол уклона русла реки (для котловины озера  $\alpha=0$ ).

Уравнение для описания динамики полей температуры и показателей качества воды модели W2 v.3.5 имеет вид:

$$\frac{\partial B\Phi}{\partial t} + \frac{\partial UB\Phi}{\partial x} + \frac{\partial WB\Phi}{\partial z} - \frac{\partial (BD_x \frac{\partial \Phi}{\partial x})}{\partial x} - \frac{\partial (BD_z \frac{\partial \Phi}{\partial z})}{\partial z} = q_{\Phi} B + S_{\Phi} B, \quad (6)$$

где  $\Phi$  – поперечно осредненные значения  $T_w$  или концентраций показателей качества воды;  $D_x$  и  $D_z$  – коэффициенты продольной и вертикальной дисперсии для  $\Phi$ ;  $q_\Phi$  – расходы боковых притоков или оттоков  $\Phi$ ;  $S_\Phi$  – слагаемое поперечного осредненного источника / стока  $\Phi$ .

При применении уравнений (1) – (6) для описания реки Чулышман используется ортогональная система координат  $(x, z)$ , где ось  $x$  направлена вдоль уклонов русла отдельных участков реки, определяемых на основе анализа тальвега (см. рис. 1).



Рис. 1. Схема задания уклона русла реки.

Описание метода декомпозиции русла на отдельные участки для определения соответствующих значений  $\alpha$  дано в разделе «Река», главы 3 [12]. При применении (1) –(6) для описания физических, химических и биологических процессов Телецкого озера использовалась декартова система координат, т.е. полагалось, что  $\alpha = 0$ .

Блоки гидродинамики версий 3.5 и 3.2 модели W2 практически совпадают. Гидродинамический блок версии 3.2 описан в первой части публикации, посвященной представлению ИВК «Северная Сосьва». В блоке расчета показателей качества воды модели W2 v.3.5, по сравнению с v.3.2 увеличено общее количество основных и производных переменных. В частности, группы фитопланктона и перифитона, описываемые версией 3.2, дополнены группами макрофитов и зоопланктона [12].

Начальные и граничные условия для системы уравнений (1) - (6) задаются, согласно [12], на основе данных наблюдений, содержащихся в БД ИВК, либо иным способом – в зависимости от целей конкретных численных экспериментов с моделью W2.

Система моделирования бассейна WMS [13], подробно описанная в первой части публикации, является в составе ИВК «Телецкое озеро» инструментом препроцессинговой обработки входных данных для подготовки модели топографии береговой зоны озера и поймы участка реки Чулышман, а также модели и файла батиметрии для W2 v.3.5.

Область моделирования ИВК «Телецкое озеро» включает в себя: 1 – участок реки Чулышман от г/п Балыкча до устья реки (длина участка - 9,8 км), 2 – меридиональную и широтную части Телецкого озера, включая Камгинский залив, 3 – исток реки Бии, 4 – устьевые зоны наиболее крупных боковых притоков к озеру, а также, береговые или буферные зоны Телецкого озера и участка реки Чулышман.

В соответствии с файловой структурой организации входной информации модели W2 база данных ИВК «Телецкое озеро» имеет следующие модули: 1 – батиметрия озера и участка реки Чулышман, топография береговой зоны; 2 – метеорология; 3 – гидрология; 4 – характеристики качества воды.

## 2. Технология и результаты автоматизированной генерации файла батиметрии модели W2

Файл батиметрии модели W2 в ИВК «Телецкое озеро» генерируется с помощью системы WMS. Ключевой момент при генерации файла батиметрии Телецкого озера и участка реки Чулышман заключается в необходимости построения нерегулярной триангуляционной сети (TIN) отметок высот над уровнем моря, адекватно описывающей как котловину и береговую зону Телецкого озера, так русло и пойму участка реки Чулышман. Для этого были использованы следующие наборы данных из модуля 1 БД ИВК:

- 1 - цифровая батиметрическая карта Телецкого озера при отметке уреза воды 434 мБС [2];
- 2 - цифровые модели рельефа, полученные из цифровой базы данных Национального управления океанических и атмосферных исследований (NOAA) США, размещенные в открытом доступе в Национальном геофизическом информационном центре (NGDC) [14];
- 3 - цифровая модель рельефа SRTM (Shuttle Radar Topography Mission – радарная топографическая миссия Шаттл) [15].

Данные об отметках высот и глубинах, содержащиеся в перечисленных выше источниках, с помощью ArcGIS 9.3 были конвертированы в единую TIN. После создания TIN производится последовательная декомпозиция моделируемого водного объекта на: 1 - крупные водные тела (WB) с характерными особенностями батиметрии, параметров экосистемы и с индивидуальным описанием метеорологического форсинга; 2 - отдельные участки или водотоки (Br) в составе WB; 3 - сегменты, на которые делятся Br; 4 - слои по z, на которые в свою очередь делятся сегменты. Основой для указанной декомпозиции при разработке ИВК «Телецкое озеро» послужила хорошо обоснованная идея авторов [1] о целесообразности разбиения акватории Телецкого озера с учетом разнообразия глубин и термического режима на пять участков (см. раздел 5.3 главы 5 [1]). Далее на основании данных высотных отметок поверхностей и глубин, содержащихся в TIN, и декомпозиции водного объекта (водные тела → водотоки → сегменты → слои) с помощью WMS генерируется файл батиметрии модели W2.

На рис. 2 показан пример расчетной сетки модели W2, сгенерированной с помощью ИВК «Телецкое озеро» и состоящей из 4 водных тел (WB1 – устьевой участок реки Чулышман, WB2 – южная оконечность озера, включая Кыгинский залив, WB3 – меридиональная часть озера, WB4 – зональная часть озера, включая Камгинский залив), 7 водотоков (Br 1-3 в составе WB1, Br4 в составе WB2, Br5 в составе WB3 и Br6-7 в составе WB4) при общем числе сегментов, равном 308. Толщины слоев воды речного участка заданы равными 0,5 м и самого озера - равными 1 м в слое от поверхности до глубины 135 м и 3 м ниже этой глубины до дна водоема. Длины сегментов в зависимости от распределения уклонов котловины озера по тальвегу изменялись от 20 до 2087 м.

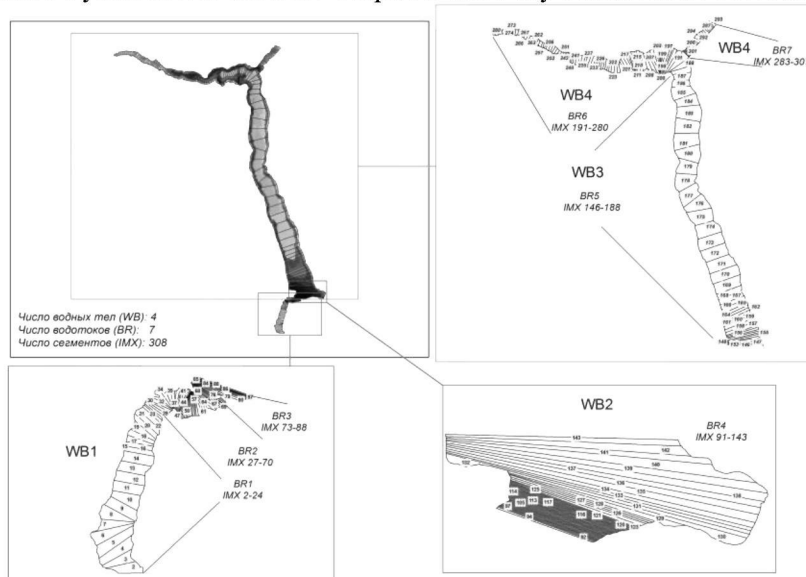


Рис. 2. Расчетная сетка устьевых участка реки Чулышман и Телецкого озера в плане.

### 3. Проблемы калибровки и верификации ИВК «Телецкое озеро». Результаты воспроизведения стадий формирования речного термического бара

В данном разделе мы кратко обсудим основные проблемы, которые необходимо преодолеть при калибровке и верификации блока гидродинамики и термического режима модели W2 v.3.5 и конкретные результаты численного воспроизведения четырех фаз весенней циркуляции, индуцируемой речным притоком в приустьевой части глубокого озера [16]. В числе упомянутых проблем остановимся на двух. Первая – это оценка адекватности и достаточности данных наблюдений, используемых при калибровке для задания начальных и граничных условий. Наиболее обременяющим для калибровки водного баланса Телецкого озера является дефицит данных гидрологических наблюдений в устьевых зонах боковых притоков к озеру (см. табл. 5 [1]). В БД ИВК «Телецкое озеро» включены имеющиеся данные наблюдений с гидропостов, расположенных на реках Чири, Кокши и Чеченег. В сумме, это около 5 % боковой приточности, оцениваемой по площади водосбора указанных притоков. Однако данные о расходах и температуре воды на реках Кыга, Боль-

шие и Малые Чили, Камга, Колдор, Самыш и др. (в сумме они составляют около 24 % боковой приточности) отсутствуют. Различные способы решения данной проблемы предлагаются в [1, 12].

По рекомендации В.В. Селегея в качестве основного года для проведения калибровки гидродинамики и термического режима озера был выбран 1968 год. Именно этот год характеризуется наиболее полным набором доступных данных гидрометеорологических наблюдений Озерной станции Западно-Сибирского УГМС, которые детально проанализированы в [1-3] и частично использовались для верификации моделей [5, 6 и др.]. Дополнительно в модули 2 и 3 БД ИВК «Телецкое озеро» были включены осредненные (среднесуточные, пентадные и декадные) гидрометеорологические данные с метеостанций и постов Озерной станции за период 1968-1972 гг., которые будут использованы для оценок внутрисезонной и межгодовой изменчивости термического режима Телецкого озера при различных сценариях боковой приточности и погодно-климатического форсинга.

Вторую проблему калибровки и верификации модели W2 для воспроизведения гидродинамики и термического режима Телецкого озера точно определяет позиция авторов [12, глава 3, стр. 20]: «Глубокое понимание процессов, которые предстоит моделировать, так же, как знание воспроизводимой системы, являются абсолютно необходимыми, если мы хотим достичь цели, на которую направлено моделирование». В нашем случае речь идет о выборе базовой феноменологической модели сезонной динамики циркуляции и термического режима Телецкого озера. Основа феноменологической модели внутригодовой эволюции термического режима озера разработана В.В. Селегеем [1-3]. В ней автор выделил 4 периода (лимносезона): 1 – период весеннего нагревания, с формированием термических баров на мелководном северо-западном участке и юге меридиональной части озера; 2 – период летнего нагревания, с различными типами вертикальной стратификации и влияния ветрового форсинга; 3 – период осеннего охлаждения с двумя последовательными фазами: в условиях прямой стратификации до момента возникновения термических баров (с реализацией механизмов перемешивания водных масс всей толщи озера при штормах) и при наличии термических баров; 4 - период зимнего минимума.

Впечатляющие результаты экспериментальных исследований и построения, на основе данных измерений высокого пространственно-временного разрешения, феноменологических моделей (или схем) сезонной эволюции механизмов циркуляции и характеристик термического режима глубоких межгорных озер Камлупс и Коотеней (Британская Колумбия, Канада), представлены в работах Кармака и соавторов [16-18]. Озера Камлупс и Коотеней по многим физико-географическим и термодинамическим характеристикам можно считать аналогами Телецкого озера.

Феноменологические модели Селегея [1] и Кармака [16-18] во многом идентичны относительно обобщенных характеристик термического режима и доминантных механизмов и типов циркуляции озер в периоды летнего нагревания, осеннего охлаждения и в зимний период. Однако существенные различия этих концептуальных моделей фиксируются для периода весеннего нагревания при анализе различных типов циркуляции, индуцируемых проникновением в озеро речных вод основного притока (река Томпсон в озере Камлупс, реки Коотеней и Дункан в озере Коотеней и река Чулышман в Телецком озере). Период весеннего нагревания в феноменологической модели Кармака [17] делится на три фазы: А – «ранняя весна», когда прогреваемые до 4 °С (температура максимальной плотности) речные воды основного притока опускаются ко дну озера вдоль склона котловины водоема, примыкающего к устью реки, образуя плотностную склоновую струю; В – «середина весны», когда дальнейшее нагревание речных вод выше 4 °С приводит к формированию речного термического бара [16] вблизи границы «озеро-река» с продолжающимся опусканием «тяжелой» смеси речной и озерной воды ко дну водоема вдоль приустьевых склонов; С – «поздняя весна», когда температура верхних слоев озера достигает значений равных или выше 4 °С, каббеллинг-процесс [16], поддерживающий речной термический бар, прекращается, а теплая речная вода вблизи устья реки деблокируется и начинает распространяться в поверхностных слоях озера, приводя к формированию первичного термоклина.

Причина отсутствия описания фаз А и В периода весеннего нагревания в феноменологической модели Селегея на участке озера, непосредственно примыкающего к устью реки Чулышман, как нам представляется, достаточно проста. Основные термодинамические события фаз А (плотностная склоновая струя) и В (речной термический бар, запирающий приток речной воды в озеро) «разыгрываются», согласно феноменологической модели Кармака, над крутым склоном котловины, примыкающим к дельте реки Томпсон на участке длиной менее 1 км (см. рис. 6 из статьи [16]). То же самое должно быть справедливо для южной оконечности Телецкого озера. Однако ближайший

озерный гидрологический разрез, на котором фиксируются вертикальные профили  $T_w$  (м. Кырсай) располагается на расстоянии около 2 км от устья Чулышмана. Таким образом, отсутствие фаз А и В, в модели Селегея, связано с отсутствием пунктов наблюдений «в нужном месте и в нужное время».

Для того чтобы проверить высказанные нами предположения, было принято решение сделать попытку воспроизвести с помощью ИВК «Телецкое озеро» схему смены 4 типов циркуляции, обусловленных проникновением в озеро быстро прогреваемых в весенний период речных вод (схему Кармака [16]), когда температура воды озера гипотетически считается величиной постоянной и равной 2 °С, а поступающая речная вода быстро прогревается от 0 до 8 °С. В схеме Кармака [16] также были проигнорированы эффекты поверхностного теплообмена, усложняющие термическую структуру открытого озера.

Всего были проведены 4 численных эксперимента (а, b, c и d), соответствующие 4 типам циркуляции схемы Кармака, которые показаны на рис. 3 [16]. По полной аналогии со схемой Кармака, в модели W2 значения температуры воды речного притока -  $T_R$  (в нашем случае  $T_w$  на г/п Балыкча) задавались равными 0, 4, 6 и 8 °С, для воспроизведения 4 типов циркуляции схемы Кармака, соответственно. Расходы воды на входном створе ( $Q_R$ ) задавались с помощью БД ИВК, как средние значения на г/п Балыкча за период 1968-72 гг. Значения  $Q_R$  были заданы равными: 20 м<sup>3</sup>/с для циркуляции типа а (30 марта среднего года); 60 м<sup>3</sup>/с для циркуляции типа b (20 апреля); 590 м<sup>3</sup>/с для циркуляции типа с (16 мая); и 950 м<sup>3</sup>/с для циркуляции типа d (26 мая). Блок теплообмена на поверхности воды в модели W2, как и в схеме Кармака, был отключен. Остальные параметры модели W2 были заимствованы из работы [9].

Результаты расчетов численных экспериментов а, b, c и d, показаны на рис. 3 в виде полей  $T_w$  и векторного представления течений. Последовательное сопоставление фрагментов а, b, c и d рис. 3 настоящей статьи и соответствующих фрагментов рис. 3 из работы [16] позволяет сделать однозначный вывод о полном феноменологическом подобии типов циркуляции в схеме Кармака и воспроизводимых с помощью ИВК «Телецкое озеро».

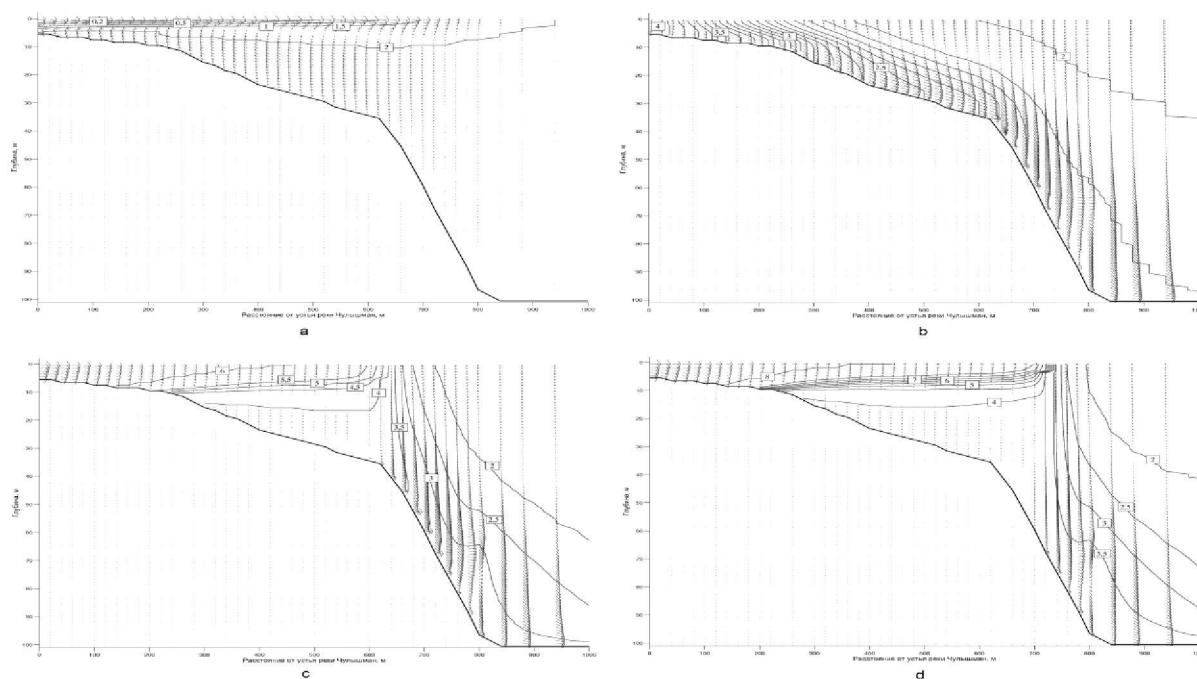


Рис. 4. Поля температуры воды и векторные отображения течений, полученные с помощью ИВК «Телецкое озеро» для сравнения с типами циркуляции схемы Кармака [16]

### Заключение

1. Наличие в научной литературе вполне адекватных феноменологических моделей циркуляции и термического режима глубоких горных озер и разработка информационно-вычислительного комплекса, включающего в себя программные средства WMS, CE-QUAL-W2 v.3.5 и соответствующую БД, позволили получить многоцелевой компьютерный инструмент, предназначенный для решения комплексных научных и практических задач, связанных с динамикой экосистемы Телец-

кого озера, о чем свидетельствует анализ серий численных экспериментов, проведенных авторами статьи, с блоком гидродинамики и теплопереноса модели W2.

2. Информационно-вычислительный комплекс также может быть использован при планировании программ будущих экспериментальных исследований Телецкого озера, например, при проведении экспедиционных работ, нацеленных на более полное изучение динамики речного и озерного термических баров.

### Благодарности

Авторы выражают признательность чл.-кор. РАН В.Н. Лыкосову, д-ру геогр. наук В.А. Земцову, д-ру физ.-мат. наук В.Н. Крупчатникову и канд. биол. наук В.В. Кириллову за поддержку данной работы, ценные советы и рекомендации. Особой благодарности заслуживает известный исследователь Телецкого озера В.В. Селегей, многочисленные обсуждения и дискуссии с которым позволили нам лучше понять, что и как надо делать.

### Литература

1. Селегей В.В., Селегей Т.С. Телецкое озеро. Гидрометеорологический режим озер и водохранилищ СССР. Л.: Гидрометеоиздат, 1978. - 143 с.
2. Селегей В., Дехандсхюттер Б., Клеркс Я., Высоцкий Е. Физико-географическая и геологическая характеристика Телецкого озера. Тервюрен – Бельгия, 2001. – 322 с.
3. Селегей В.В. Телецкое озеро: очерки истории: в 3 кн. Барнаул: Пять плюс, 2011. – Кн. 3. – 244 с.
4. Васильев О.Ф., В.В. Селегей, В.В. Кириллов. Озеро Телецкое – особенности лимнологической системы: история и перспективы исследования // Байкал – природная лаборатория для исследования изменения окружающей Среды и климата. Иркутск: Лисна, 1994. Т. 6. 50 с.
5. Vasiliev O.F., Kvon V.I., Kvon D.V. Mathematical modeling of thermal bar in Lake Teletskoye. Proceedings of the 28 IAHR Congress, Theme D, Graz (Austria), 1999.6 p.
6. Квон Д.В., Квон В.И., Семчуков А.Н. Численный расчет продольно-вертикальной термической структуры Телецкого озера в годовом цикле // Вычислительные технологии. 2000. Т. 5, № 3. С. 29-45.
7. Черных Д.В., Ротанова И.Н. Современные ландшафты Прителецкого района (Горный Алтай) // География и природные ресурсы. 2006. № 1. С. 150-154.
8. Daniel P. Loucks and Eelco van Beek with contributions from Jerry R. Stedinger, Jozef P.M. Dijkman, Monique T. Villars. Water Resources Systems Planning and Management An Introduction to Methods, Models and Applications. Studies and Reports in Hydrology. UNESCO PUBLISHING. – 2005. – 676 p.
9. Berger C.J. and Wells S.A. Lake Whatcom Water Quality Model. Technical Report EWR-03-05, Maseeh College of Engineering and Computer Science, Department of Civil and Environmental Engineering, Portland State University, Portland, OR, 2005. - 839 p.
10. Lake Whatcom Models Review. Independent review performed by Tetra Tech for the Washington State Department of Ecology, April 2008 Publication no. 09-10-013. 38 p.
11. Danchev V. N., Pushistov P. Yu., Kirillov V. V., Dyachenko A. V. Information system, as the tool of research of an ecosystem dynamics of a deep reservoir and its basin (on example of lake Teletskoye) / International conference on Environmental Observations, Modeling and Information Systems ENVIROMIS-2010, July 5-11, Tomsk, Russia: Program&Abstracts, P 112-113
12. Cole T. M. and Wells S. A. "CE-QUAL-W2" A two-dimensional, laterally averaged, Hydrodynamic and Water Quality Model, Version 3.5, Instruction Report EL-06-1, US Army Engineering and Research Development Center, Vicksburg, MS, 2006. – 681 p.
13. URL: <http://www.aquaveo.com>
14. URL: <http://www.ngdc.noaa.gov/>
15. URL: <http://srtm.csi.cgiar.org/>
16. Carmack, E. C. Combined influence of inflow and lake temperatures on spring circulation in a riverine lake, J. Phys. Oceanogr. 9, 1978. P. 422–434.
17. Carmack, E. C., Gray, C. B. J., Pharo, C. H. & Daley, R. J. Importance of lake-river interaction on seasonal patterns in the general circulation of Kamloops Lake, British Columbia. Limnol. Oceanogr. 24(4), 1979. P. 634–644.
18. Carmack, E. C., Wiegand, R. C., Daley, R. J., Gray, C. B. J., Jasper, S. & Pharo, C. H. Mechanisms influencing the circulation and distribution of water mass in a medium residence time lake. Limnol. Oceanogr. 31(2), 1986. P. 249–265.

*В.Н. Данчев, П.Ю. Пушистов. Опыт разработки и результаты применения информационно-вычислительного комплекса для моделирования гидродинамики и качества воды рек и водоемов бассейна Оби.*

---

*Пушистов Петр Юрьевич, д-р физ.-мат. наук, проф., научный консультант, Институт водных и экологических проблем Сибирского отделения Российской академии наук, Россия, 656038, г. Барнаул, ул. Молодежная, 1; e-mail: [pushtcmpr@mail.ru](mailto:pushtcmpr@mail.ru)*

*Данчев Василий Николаевич, аспирант, Институт водных и экологических проблем Сибирского отделения Российской академии наук, Россия, 656038, г. Барнаул, ул. Молодежная, 1; e-mail: [danchev\\_vasilii@mail.ru](mailto:danchev_vasilii@mail.ru)*

*Pushistov Peter Yurievich, doctor of physical and mathematical sciences, professor, scientific adviser, Institute for Water and Environmental Problems Siberian Branch of the RAS, Russia, 656038, Barnaul, ul. Molodyoznaya, 1.*

*Danchev Vasily Nikolayevich, postgraduate, Institute for Water and Environmental Problems Siberian Branch of the RAS, Russia, 656038, Barnaul, ul. Molodyoznaya, 1.*

УДК 537.876:537.622.4

©Г.Б. Итигилов, Д.Ш. Ширапов

## МЕТОД ИНВАРИАНТНЫХ ПРЕОБРАЗОВАНИЙ ДЛЯ ОПРЕДЕЛЕНИЯ ПОПЕРЕЧНЫХ КОМПОНЕНТ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В ГИРОТРОПНЫХ ОГРАНИЧЕННЫХ ОБЛАСТЯХ

Методом инвариантных преобразований получены аналитические выражения для поперечных компонент электромагнитного поля в закрытой гиротропной области с ортогональной криволинейной формой поперечного сечения при произвольном намагничивании. Такой подход позволяет легко перейти к конкретным формам поперечного сечения (прямоугольной, круглой и эллиптической) при различных случаях намагничивания (продольном, касательном, нормальном). В данной работе рассматривается продольное намагничивание.

**Ключевые слова:** поперечные компоненты электромагнитного поля, намагничивание, гиротропное заполнение, ограниченная область, символы Кристоффеля, коэффициенты Ламэ.

*G.B. Itigilov, D.Sh. Shirapov*

## METHOD OF INVARIANT TRANSFORMATIONS FOR DEFINITION CROSS-SECTION A COMPONENT OF AN ELECTROMAGNETIC FIELD IN GIROTROPIC THE LIMITED AREAS

The method of invariant transformations receives analytical expressions for cross-section a component of an electromagnetic field in closed girotropic areas with the orthogonal curvilinear form of cross-section section at any magnetization. Such approach allows to proceed easily to concrete forms of cross-section section (rectangular, round and elliptic) at various cases of magnetization (longitudinal, касательном, normal). In the given work longitudinal magnetization is considered.

**Keywords:** cross-section components of an electromagnetic field, magnetization, girotropic filling, limited area, symbols of Kristoffel, coefficients of the Lame.

### Введение

Распространение электромагнитных волн (ЭМВ) в намагниченной ферритовой (гиротропной) среде характеризуется тем, что фазовая скорость, затухание и поляризация распространяющейся волны зависят от величины напряженности внешнего магнитного поля и его направления относительно направления распространения волны. Вследствие этого условия распространение волн в направляющих системах с гиротропным заполнением можно сознательно изменять в широких пределах, изменяя величину и направление магнитного поля [1].

Для исследования условий распространения электромагнитных волн в регулярной гиротропной ограниченной области с криволинейной ортогональной формой поперечного сечения, намагниченной вдоль одной из координатных осей, необходимы инвариантные преобразования на основе тензорного исчисления (метод инвариантных преобразований – МИП). Удобство применения МИП для математического анализа ограниченных областей с обобщенно-ортогональной формой поперечного сечения является то, что метод обладает свойством инвариантности относительно преобразования координат.

В общем случае рассматривается намагничивание феррита вдоль одной из трех координатных осей [2]. При этом рассматривают три случая кривизны поперечных координат:

- 1) линейность по обеим координатным осям;
- 2) кривизна по одной из координатных осей;
- 3) кривизна по обеим координатным осям.

Первому случаю соответствует прямоугольная система координат, второму – цилиндрическая, третьему – эллиптическая.

Целью данной статьи является получение аналитических выражений поперечных компонент электромагнитного поля (ЭМП) в гиротропной ограниченной области при произвольном и продольном намагничиваниях с использованием МИП, описанного в §1 и 2.



### 1. Характеристики ортогональных систем координат

Определим характеристики эллиптической системы координат: коэффициенты Ламэ, символы Кристоффеля и метрику.

Коэффициенты Ламэ определяем по формуле [3]:

$$h_i = \sqrt{\left(\frac{\partial x}{\partial u_i}\right)^2 + \left(\frac{\partial y}{\partial u_i}\right)^2 + \left(\frac{\partial z}{\partial u_i}\right)^2}, \quad (1)$$

где  $u_1 = \xi, u_2 = \varphi, u_3 = z$  – эллиптические координаты.

Тогда из (1) имеем:

$$h_1 = h_2 = ed; \quad h_3 = 1, \quad (2)$$

где  $d = \sqrt{ch^2\xi - \cos^2\varphi}$ ,  $e$  – фокусное расстояние.

Символы Кристоффеля определяем из [4]. После преобразований:

$$\begin{cases} \Gamma_{ij}^1 = \frac{1}{2d^2} \begin{bmatrix} sh2\xi & \sin 2\varphi & 0 \\ \sin 2\varphi & -sh2\xi & 0 \\ 0 & 0 & 0 \end{bmatrix}; \\ \Gamma_{ij}^2 = \frac{1}{h_2} \begin{bmatrix} -\sin 2\varphi & sh2\xi & 0 \\ sh2\xi & \sin 2\varphi & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{cases} \quad (3)$$

Тогда метрика будет равна:

$$ds^2 = (h_1 du_1)^2 + (h_2 du_2)^2 + (h_3 du_3)^2 = e^2 d^2 d\xi^2 + e^2 d^2 d\varphi^2 + dz^2. \quad (4)$$

Аналогично для цилиндрической системы координат ( $u_1 = r, u_2 = \phi, u_3 = z$ ) коэффициенты Ламэ определим из [3], а символы Кристоффеля из [4]. После преобразований:

$$\begin{cases} h_1 = h_3 = 1, \quad h_2 = r; \\ \Gamma_{11}^1 = \Gamma_{22}^2 = \Gamma_{12}^1 = 0, \quad \Gamma_{21}^2 = \frac{1}{r}. \end{cases} \quad (5)$$

Метрика будет равна:

$$ds^2 = (h_1 du_1)^2 + (h_2 du_2)^2 + (h_3 du_3)^2 = dr^2 + r^2 d\phi^2 + dz^2. \quad (6)$$

Для декартовой системы координат ( $u_1 = x, u_2 = y, u_3 = z$ ):

$$h_1 = h_2 = h_3 = 1, \quad \Gamma_{11}^1 = \Gamma_{22}^2 = \Gamma_{12}^1 = \Gamma_{21}^2 = 0. \quad (7)$$

Метрика будет равна:

$$ds^2 = (h_1 du_1)^2 + (h_2 du_2)^2 + (h_3 du_3)^2 = dx^2 + dy^2 + dz^2. \quad (8)$$

### 2. Поперечные компоненты ЭМП

При рассмотрении процессов, гармонических во времени (зависимость от времени примем в виде  $e^{j\omega t}$ ), уравнения Максвелла без наведенных токов и зарядов имеют вид [1]:

$$\begin{cases} rot H = j\omega E; \quad rot E = -j\omega B; \\ div D = 0; \quad div B = 0, \end{cases} \quad (9)$$

где  $E, H$  – соответственно напряженности электрического и магнитного полей;  $D, B$  – соответственно электрическая и магнитная индукции,  $j$  – мнимая единица,  $\omega$  – циклическая частота.

Система (9) дополняется материальными уравнениями среды:

$$D = \epsilon E; \quad B = \|\mu\| H, \quad (10)$$

где  $\epsilon$  – абсолютная диэлектрическая проницаемость, а  $\|\mu\|$  – тензор магнитной проницаемости.

При произвольном намагничивании, когда внешнее намагничивающее постоянное поле имеет составляющие по всем трем координатным осям, тензор магнитной проницаемости феррита, как следует из [5], имеет вид:

$$\mu_{ik} = \begin{bmatrix} \mu_{11} & jk & jl \\ -jk & \mu_{22} & jm \\ -jl & -jm & \mu_{33} \end{bmatrix}, \quad (11)$$

где  $\mu_{11}, \mu_{22}, \mu_{33}, k, l, m$  – компоненты тензора.

В системе (9), разложив  $rot H$  и  $rot E$  по осям, после подстановок и преобразований получим поперечные компоненты электромагнитного поля (ЭМП) для гиротропной области с криволинейной ортогональной формой поперечного сечения:

$$\left\{ \begin{array}{l} E_1 = -\frac{j\gamma a^2}{a^2 b^2 - w^4 \varepsilon^2 k^2} \left\{ \begin{array}{l} \nabla_1 E_3 + \left( \frac{w\mu_{22}}{\gamma} \frac{p^2}{a^2} \nabla_2 + wm \right) H_3 - \\ - \frac{jw^2 \varepsilon k}{a^2} \left[ \nabla_2 E_3 - \left( \frac{\gamma}{w\varepsilon} \nabla_1 + wl \right) H_3 \right] \end{array} \right\}, \\ E_2 = -\frac{j\gamma b^2}{a^2 b^2 - w^4 \varepsilon^2 k^2} \left\{ \begin{array}{l} \nabla_2 E_3 - \left( \frac{w\mu_{11}}{\gamma} \frac{g^2}{b^2} \nabla_1 + wl \right) H_3 + \\ + \frac{jw^2 \varepsilon k}{b^2} \left[ \nabla_1 E_3 + \left( \frac{\gamma}{w\varepsilon} \nabla_2 + wm \right) H_3 \right] \end{array} \right\}, \\ H_1 = \frac{j\gamma b^2}{a^2 b^2 - w^4 \varepsilon^2 k^2} \left\{ \begin{array}{l} \frac{w\varepsilon}{\gamma} \nabla_2 E_3 - \left( \nabla_1 + \frac{w^2 \varepsilon l}{\gamma} \right) H_3 + \\ + \frac{jw^2 k \varepsilon}{b^2} \left[ \frac{w\varepsilon}{\gamma} \nabla_1 E_3 + \left( \nabla_2 + \frac{w^2 \varepsilon m}{\gamma} \right) H_3 \right] \end{array} \right\}, \\ H_2 = -\frac{j\gamma a^2}{a^2 b^2 - w^4 \varepsilon^2 k^2} \left\{ \begin{array}{l} \frac{w\varepsilon}{\gamma} \nabla_1 E_3 + \left( \nabla_2 + \frac{w^2 \varepsilon m}{\gamma} \right) H_3 - \\ - \frac{jw^2 k \varepsilon}{a^2} \left[ \frac{w\varepsilon}{\gamma} \nabla_2 E_3 - \left( \nabla_1 + \frac{w^2 \varepsilon l}{\gamma} \right) H_3 \right] \end{array} \right\}, \end{array} \right. \quad (12)$$

где

$$a^2 = w^2 \mu_{11} \varepsilon - \gamma^2; \quad b^2 = w^2 \mu_{22} \varepsilon - \gamma^2; \quad g^2 = w^2 \varepsilon \frac{\mu_{11} \mu_{22} - k^2}{\mu_{11}} - \gamma^2;$$

$$p^2 = w^2 \varepsilon \frac{\mu_{11} \mu_{22} - k^2}{\mu_{22}} - \gamma^2; \quad \nabla_1 = \frac{1}{h_1} \frac{\partial}{\partial u_1}; \quad \nabla_2 = \frac{1}{h_2} \frac{\partial}{\partial u_2},$$

$w$  – циклическая частота,  $\gamma$  – постоянная распространения.

Выражение (12) описывает поперечные компоненты ЭМП в ограниченной гиротропной области с ортогональной формой поперечного сечения при произвольном намагничивании.

При продольном намагничивании тензор магнитной проницаемости феррита имеет вид [5]:

$$\overset{\leftrightarrow}{\mu} = \begin{bmatrix} \mu & jk & 0 \\ -jk & \mu & 0 \\ 0 & 0 & \mu_{\parallel} \end{bmatrix}, \quad (13)$$

где  $\frac{\mu}{\mu_0} = 1 + \frac{w_0 w_M}{w_0^2 - w^2}$ ;  $\frac{k}{\mu_0} = \frac{w w_M}{w_0^2 - w^2}$ ;  $w_M = \mu_0 Y M_0$ ;  $Y = 1.76 * 10^{11} \frac{Kл}{кг}$  – гиромангнитное отношение

для спина;  $w_0 = \mu_0 Y H_0$  – частота ферромагнитного резонанса;  $M_0$  – намагниченность насыщения феррита;  $H_0$  – внешнее намагничивающее магнитное поле.

Из формулы (12), учитывая (13), получим поперечные компоненты ЭМП в ограниченной гиротропной области с криволинейной ортогональной формой поперечного сечения при продольном намагничивании:

$$\begin{cases} E_1 = -\frac{j\gamma a^2}{g_+ g_-} \left\{ \nabla_1 E_3 + \frac{w\mu c^2}{\gamma a^2} \nabla_2 H_3 - \frac{jw^2 \varepsilon k}{a^2} \left[ \nabla_2 E_3 - \frac{\gamma}{w\varepsilon} \nabla_1 H_3 \right] \right\}, \\ E_2 = -\frac{j\gamma a^2}{g_+ g_-} \left\{ \nabla_2 E_3 - \frac{w\mu c^2}{\gamma a^2} \nabla_1 H_3 + \frac{jw^2 \varepsilon k}{a^2} \left[ \nabla_1 E_3 + \frac{\gamma}{w\varepsilon} \nabla_2 H_3 \right] \right\}, \\ H_1 = \frac{j\gamma a^2}{g_+ g_-} \left\{ \frac{w\varepsilon}{\gamma} \nabla_2 E_3 - \nabla_1 H_3 + \frac{jw^2 \varepsilon k}{a^2} \left[ \frac{w\varepsilon}{\gamma} \nabla_1 E_3 + \nabla_2 H_3 \right] \right\}, \\ H_2 = -\frac{j\gamma a^2}{g_+ g_-} \left\{ \frac{w\varepsilon}{\gamma} \nabla_1 E_3 + \nabla_2 H_3 - \frac{jw^2 \varepsilon k}{a^2} \left[ \frac{w\varepsilon}{\gamma} \nabla_2 E_3 - \nabla_1 H_3 \right] \right\}, \end{cases} \quad (14)$$

где

$$a^2 = w^2 \mu_1 \varepsilon - \gamma^2 = w^2 \mu \varepsilon - \gamma^2; \quad g_{\pm}^2 = w^2 \varepsilon \mu \pm w^2 \varepsilon k - \gamma^2;$$

$$c^2 = w^2 \varepsilon \frac{\mu^2 - k^2}{\mu} - \gamma^2.$$

Из (14) можно легко получить поперечные компоненты ЭМП для конкретных форм ограниченных областей. При этом форма поперечного сечения ограниченной области определяет выбор системы координат. При анализе ЭМВ в ограниченных областях с эллиптической, круглой и прямоугольной формами поперечного сечения, используются соответственно эллиптическая, цилиндрическая и декартова системы координат.

Чтобы определить поперечные компоненты ЭМП гиротропной эллиптической области, в (14) необходимо подставить (2):

$$\begin{cases} E_{\xi} = -\frac{j\gamma a^2}{g_+ g_-} \frac{1}{ed} \left[ \frac{\partial E_z}{\partial \xi} + \frac{w\mu c^2}{\gamma a^2} \frac{\partial H_z}{\partial \varphi} - \frac{jw^2 \varepsilon k}{a^2} \left( \frac{\partial E_z}{\partial \varphi} - \frac{\gamma}{w\varepsilon} \frac{\partial H_z}{\partial \xi} \right) \right], \\ E_{\varphi} = -\frac{j\gamma a^2}{g_+ g_-} \frac{1}{ed} \left[ \frac{\partial E_z}{\partial \varphi} - \frac{w\mu c^2}{\gamma a^2} \frac{\partial H_z}{\partial \xi} + \frac{jw^2 \varepsilon k}{a^2} \left( \frac{\partial E_z}{\partial \xi} + \frac{\gamma}{w\varepsilon} \frac{\partial H_z}{\partial \varphi} \right) \right], \\ H_{\xi} = \frac{j\gamma a^2}{g_+ g_-} \frac{1}{ed} \left[ \frac{w\varepsilon}{\gamma} \frac{\partial E_z}{\partial \varphi} - \frac{\partial H_z}{\partial \xi} + \frac{jw^2 \varepsilon k}{a^2} \left( \frac{w\varepsilon}{\gamma} \frac{\partial E_z}{\partial \xi} + \frac{\partial H_z}{\partial \varphi} \right) \right], \\ H_{\varphi} = -\frac{j\gamma a^2}{g_+ g_-} \frac{1}{ed} \left[ \frac{w\varepsilon}{\gamma} \frac{\partial E_z}{\partial \xi} + \frac{\partial H_z}{\partial \varphi} - \frac{jw^2 \varepsilon k}{a^2} \left( \frac{w\varepsilon}{\gamma} \frac{\partial E_z}{\partial \varphi} - \frac{\partial H_z}{\partial \xi} \right) \right]. \end{cases} \quad (15)$$

Аналогично из (14) с учетом (5) получим выражения для поперечных компонент ЭМП для гиротропной цилиндрической области при продольном намагничивании:

$$\begin{cases} E_r = -\frac{j\gamma a^2}{g_+ g_-} \left[ \frac{\partial E_z}{\partial r} + \frac{w\mu c^2}{\gamma a^2} \frac{1}{r} \frac{\partial H_z}{\partial \varphi} - \frac{jw^2 \varepsilon k}{a^2} \left( \frac{1}{r} \frac{\partial E_z}{\partial \varphi} - \frac{\gamma}{w\varepsilon} \frac{\partial H_z}{\partial r} \right) \right], \\ E_{\varphi} = -\frac{j\gamma a^2}{g_+ g_-} \left[ \frac{1}{r} \frac{\partial E_z}{\partial \varphi} - \frac{w\mu c^2}{\gamma a^2} \frac{\partial H_z}{\partial r} + \frac{jw^2 \varepsilon k}{a^2} \left( \frac{\partial E_z}{\partial r} + \frac{\gamma}{w\varepsilon} \frac{1}{r} \frac{\partial H_z}{\partial \varphi} \right) \right], \\ H_r = \frac{j\gamma a^2}{g_+ g_-} \left[ \frac{w\varepsilon}{\gamma} \frac{1}{r} \frac{\partial E_z}{\partial \varphi} - \frac{\partial H_z}{\partial r} + \frac{jw^2 \varepsilon k}{a^2} \left( \frac{w\varepsilon}{\gamma} \frac{\partial E_z}{\partial r} + \frac{1}{r} \frac{\partial H_z}{\partial \varphi} \right) \right], \\ H_{\varphi} = -\frac{j\gamma a^2}{g_+ g_-} \left[ \frac{w\varepsilon}{\gamma} \frac{\partial E_z}{\partial r} + \frac{1}{r} \frac{\partial H_z}{\partial \varphi} - \frac{jw^2 \varepsilon k}{a^2} \left( \frac{w\varepsilon}{\gamma} \frac{1}{r} \frac{\partial E_z}{\partial \varphi} - \frac{\partial H_z}{\partial r} \right) \right]. \end{cases} \quad (16)$$

Также аналогично из (14) с учетом (7) получим выражения для поперечных компонент ЭМП для гиротропной прямоугольной области при продольном намагничивании:

$$\begin{cases} E_x = -\frac{j\gamma a^2}{g_+ g_-} \left[ \frac{\partial E_z}{\partial x} + \frac{w\mu c^2}{\gamma a^2} \frac{\partial H_z}{\partial y} - \frac{jw^2 \epsilon k}{a^2} \left( \frac{\partial E_z}{\partial y} - \frac{\gamma}{w\epsilon} \frac{\partial H_z}{\partial x} \right) \right], \\ E_y = -\frac{j\gamma a^2}{g_+ g_-} \left[ \frac{\partial E_z}{\partial y} - \frac{w\mu c^2}{\gamma a^2} \frac{\partial H_z}{\partial x} + \frac{jw^2 \epsilon k}{a^2} \left( \frac{\partial E_z}{\partial x} + \frac{\gamma}{w\epsilon} \frac{\partial H_z}{\partial y} \right) \right], \\ H_x = \frac{j\gamma a^2}{g_+ g_-} \left[ \frac{w\epsilon}{\gamma} \frac{\partial E_z}{\partial y} - \frac{\partial H_z}{\partial x} + \frac{jw^2 \epsilon k}{a^2} \left( \frac{w\epsilon}{\gamma} \frac{\partial E_z}{\partial x} + \frac{\partial H_z}{\partial y} \right) \right], \\ H_y = -\frac{j\gamma a^2}{g_+ g_-} \left[ \frac{w\epsilon}{\gamma} \frac{\partial E_z}{\partial x} + \frac{\partial H_z}{\partial y} - \frac{jw^2 \epsilon k}{a^2} \left( \frac{w\epsilon}{\gamma} \frac{\partial E_z}{\partial y} - \frac{\partial H_z}{\partial x} \right) \right]. \end{cases} \quad (17)$$

В [6] получены поперечные компоненты ЭМП в цилиндрической и прямоугольной областях с продольно намагниченными тензорными магнитным и диэлектрическим заполнениями. Результаты, полученные в [6] полностью совпадают с выражениями (16) и (17), если среда заполнения характеризуется тензором магнитной и скалярной диэлектрической проницаемостями. Такой результат подтверждает правильность данных формул, полученных на основе инвариантных преобразований.

### Заключение

1. Применением «метода инвариантных преобразований» получена общая формула (14) для поперечных компонент ЭМП для произвольно намагниченных гиротропных ограниченных областей с обобщенно-ортогональным поперечным сечением.

2. Из (14) легко определяются соответствующие формулы поперечных компонент ЭМП для конкретных форм поперечных сечений ограниченных областей (эллиптическая, круглая, прямоугольная) для различных случаев намагничивания (продольное, касательное и нормальное).

3. В данной работе представлены переходы от поперечных компонент ЭМП для произвольно-намагниченных гиротропных ограниченных областей с обобщенно-ортогональным поперечным сечением к продольно-намагниченным гиротропным областям с эллиптической, круглой, прямоугольной формами поперечного сечения.

### Литература

1. Микаэлян А.Л. Теория и применение ферритов на сверхвысоких частотах. М.-Л.: Госэнергоиздат, 1963. 664 с.
2. Неганов В.А., Нефедов Е.И., Яровой Г.П. Современные методы проектирования линий передач и резонаторов сверх- и крайне высоких частот. М.: Педагогика-Пресс, 1998. 328 с.
3. Анго А. Математика для электро- и радиоинженеров. М.: Наука, 1967. 780 с.
4. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука, 1973. 831 с.
5. Гуревич А.Г., Мелков Г.А. Магнитные колебания и волны. М.: Физматлит, 1994. 464 с.
6. Виприцкий Д.Д. Открытые и экранированные направляющие структуры с продольно намагниченными ферритовыми слоями: дис. ... канд. техн. наук. Нижний Новгород, 2007. 177 с.

*Итигилов Гарма Борисович*, старший преподаватель кафедры электронных вычислительных систем Восточно-Сибирского государственного университета технологий и управления, тел. 8-3012-21-53-14. 670013, г. Улан-Удэ, ул. Ключевская 40В, кафедра электронных вычислительных систем, e-mail: Gablz@mail.ru  
*Ширапов Дашадондок Шагдарович*, д-р физ.-мат. наук, проф., зав. каф. электронных вычислительных систем Восточно-Сибирского государственного университета технологий и управления, тел. 8-3012-21-53-14. 670013, Республика Бурятия, г. Улан-Удэ, ул. Ключевская 40В, кафедра электронных вычислительных систем, e-mail: Shir@esstu.ru

*Itigilov Garma Borisovich*, senior teacher of chair "Electronic computer systems" of the East Siberian State University of Technology and Management.

*Shirapov Dashadondok Shagdarovich*, the doctor of physical and mathematical sciences, the professor, head of the chair "Electronic computer systems" of the East Siberian State University of Technology and Management.

УДК 51-73

© Л.А. Компаниец, Т.В. Якубайлик, О.С. Володько

## АНАЛИЗ ХАРАКТЕРИСТИК ОЗЕРА ШИРА НА ОСНОВЕ НАТУРНЫХ НАБЛЮДЕНИЙ<sup>1</sup>

В работе обобщаются результаты измерений скорости в озере Ши́ра в течение трех летних сезонов 2009-2011 гг. Описываются приборы, которыми проводились измерения, приводятся результаты натуральных измерений.

**Ключевые слова:** стратифицированное озеро, нестационарные волновые течения.

*L.A. Kompaniets, T.V. Yakubaylik, O.S. Volodko*

## ANALYSIS OF THE CHARACTERISTICS OF THE SHIRA LAKE ON THE BASIS OF FIELD OBSERVATIONS

This paper summarizes the results of measurements of the velocities in Lake Shira during three summers of 2009-2011. Devices by which measurements were carried out are described, results of natural measurements are given.

**Keywords:** the stratified lake, non-stationary wave currents.

### Введение

В настоящее время значительное влияние уделяется изучению процессов в водных экосистемах, к которым относится соленое озеро Ши́ра, расположенное в Республике Хакасия. Озеро представляет собой уникальный природный комплекс, который во многом аналогичен экологическим морским системам и обладает многими уникальными свойствами, в том числе и лечебными.

Ши́ра представляет собой бессточное озеро без островов, в которое впадает одна речка Сон. В силу малости притока все влияние реки сосредоточено в приустьевой зоне, поэтому основным внешним фактором, определяющим течение в озере, является ветровое воздействие. Озеро имеет овальную форму, длина – 9,4 км, ширина – 5 км, площадь водного зеркала – 34,7 км<sup>2</sup>, средняя глубина – 11,2 м, максимальная глубина за последние несколько десятилетий изменялась от 21 м до 24 м (по последним наблюдениям – 25 м).

В течение длительного времени в Институте биофизики СО РАН на базе научного стационара проводятся исследования функционирования экосистемы озера Ши́ра, при этом данных о характере гидрофизических процессов, таких, как скорости течений, практически не было.

Гидродинамику озера необходимо изучать в силу того, что это небольшой объект, который можно представить, как аналог для изучения других более крупных экосистем.

Как показывает опыт исследования гидрофизического режима озера Ши́ра, каждый из методов в цепочке: математическая модель, численный эксперимент, натурный эксперимент, имеет свои преимущества и недостатки одновременно.

Аналитические решения позволяют найти скорости течения в каждой точке пространства, но могут быть получены, как правило, только в очень частных случаях. Если прибавить к этому тот факт, что при нахождении аналитических решений используются очень важные, но сложно измеряемые величины (такие, как коэффициенты вертикального и горизонтального обмена), то становится понятно, что делать выводы только на основе аналитических решений необдуманно.

Численные решения в задачах гидрофизики широко применяются, но при выводе основополагающих уравнений делаются предположения, в результате которых, например, для водоемов, у которых длина и ширина значительно превышают глубину (к таковым относится озеро Ши́ра), вертикальная составляющая скорости находится из уравнения неразрывности, что влечет за собой проблемы при постановке граничных условий. Большое значение также имеет шаг сетки по пространству, а насколько его можно измельчить (а, значит, получить тонкие эффекты движения) определяется мощностью используемой машины.

---

<sup>1</sup> Работа выполнена при финансовой поддержке междисциплинарного интеграционного проекта СО РАН №56.

В настоящее время появилась возможность использовать современное оборудование для измерения вертикальной и горизонтальных компонент скорости течения, используя акустические профилографы течений (ADCP с частотой 1200 и 600 МГц), основанные на эффекте Доплера. Такие приборы используются в Институте биофизики СО РАН (ИБФ СО РАН) и Институте вычислительного моделирования СО РАН (ИВМ СО РАН). С помощью приборов ADCP можно определить возможный характер влияния перемещения частиц жидкости на распределение различных компонент экосистемы озера Шира, таких как концентрация биогенных элементов, температуры, солености, плотности, а так же для того, чтобы оценить характер распределения фитопланктона и зоопланктона под действием течений.

Акустический доплеровский профилограф течения (Acoustic Doppler Current Profiler, далее — ADCP) дает возможность применить новую технологию гидрометрических наблюдений и в настоящее время широко используется для измерения поля скоростей в поперечном сечении потока и расхода воды в реках, а также в озерных и морских водах. Он используется при проведении гидрометрических съемок, ведении мониторинга поверхностных вод и в научных исследованиях.

Казалось, что измерения скорости с помощью приборов нового поколения ADCP и SonTek способны решить эту задачу, если бы не ряд проблем.

При изменчивой и сильно нестационарной метеорологической картине в районе озера Шира в летний период (ветер меняется в зависимости от положения точки на озере, сильные волны) сам процесс измерения становится сложным, также как и интерпретация полученных данных.

Постановка натурных экспериментов является достаточно сложной задачей, в которой необходимо осуществлять координатную привязку станций, на которых проводились наблюдения, учитывать пространственное положение приборов на поверхности и дне озера. При этом необходимо принимать во внимание характер грунта на дне озера, текущие ветровые условия, которые могут оказать существенное влияния на проведения экспериментов.

Измеренные данные нуждаются в более удобном представлении для анализа, то есть в структурированном размещении в базе данных. В дальнейшем, используя эти данные, можно разработать программную среду моделирования гидрофизических процессов в трехмерном пространстве. Результат работы такой программы будет являться основой для прогноза функционирования экосистемы водоемов, озер, морей, что даст очень значительный вклад в изучение и понимание гидродинамики и гидрофизики многих водных объектов.

В данной статье систематически излагаются полученные в результате натурных наблюдений результаты и проводится их первоначальный анализ.

Сравнение с аналитическими решениями и численными расчетами будет отражено в дальнейших публикациях.

### 1. Анализ характеристик озера на основе измерения температуры и солености

Натурные измерения проводились в течение июля 2009-2010 г. на оз. Шира по нескольким направлениям.

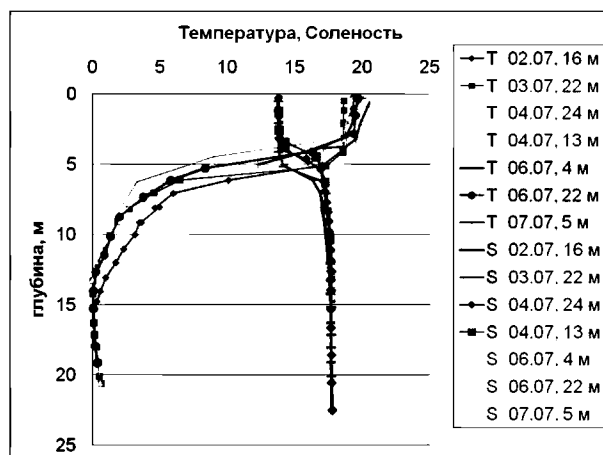


Рис. 1. Данные зонда 2010 г., июль

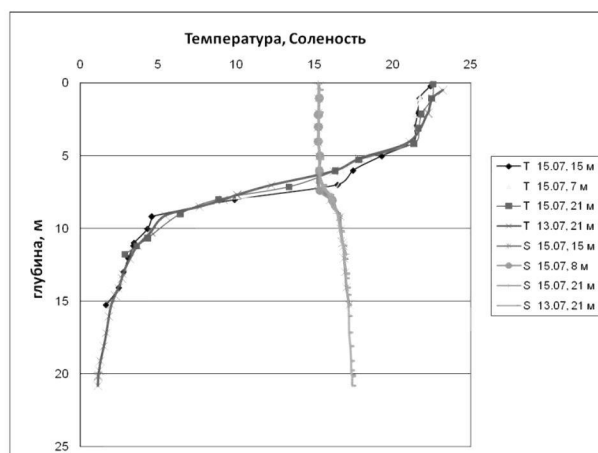


Рис. 2. Данные зонда 2009 г., июль

Точка резкого изменения градиента температуры по сравнению с 2009 г. поднялась на 1 м и располагается на глубине 4 м (было 5 м).

Соленость рассчитывалась по измеренной кондуктивности по формуле, предоставленной сотрудником Института биофизики СО РАН канд. физ.-мат. наук Д.Ю. Рогозиным:

$$S = 0,8286 \cdot C + 0,1937. \quad (1)$$

Здесь  $S$  — соленость,  $C$  — кондуктивность.

Плотность можно считать по различным формулам. Первая из них — это формула равномерно соленой воды [7], в которой плотность зависит только от температуры

$$\rho = \rho_0(1 - 1,5 \cdot 10^{-4} T). \quad (2)$$

Рисунок 3 показывает плотность, рассчитанную по этой формуле, и приращение плотности, отнесенное к  $\rho_0 = 1000 \text{ кг/м}^3$ .

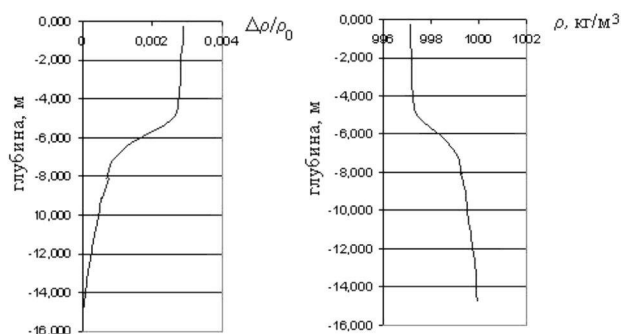


Рис. 3. Плотность, приращение плотности, формула (2)

Видно, что приращение  $\Delta\rho/\rho_0$  имеет значение 0,003.

Вторая формула — это формула Буссинеска

$$\rho(T, S) = \rho_1 \cdot \left( \varepsilon_1 + \varepsilon_2 \cdot \frac{T}{T_0} + \varepsilon_3 \cdot \frac{S}{S_0} \right), \quad (3)$$

где  $\rho_1 = 1,02541 \text{ г/см}^3$ ;  $T_0 = 17,5 \text{ }^\circ\text{C}$ ;  $S_0 = 35\text{‰}$ ;  $\varepsilon_1 = 0,97529$ ;  $\varepsilon_2 = -0,00317$ ;  $\varepsilon_3 = 0,02737$ .

На рис. 4 изображены плотность и приращение плотности в соответствии с этой формулой.

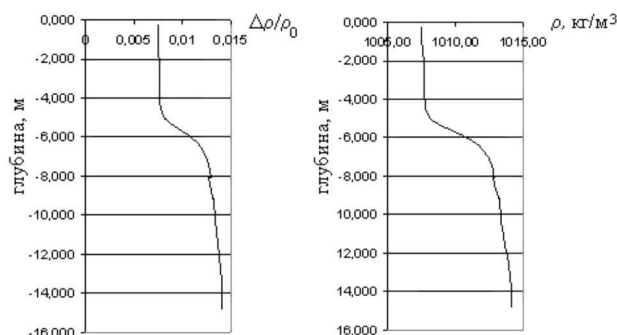


Рис. 4. Плотность, приращение плотности, формула (3)

Приращение  $\Delta\rho/\rho_0$  имеет значение 0,007.

Простой, но достаточно точной для уравнения состояния соленой воды является формула Мамаева

$$\rho(T,S) = (10^3 + [28,14 - 0,0735 \cdot T - 0,00469 \cdot T^2 + (0,802 - 0,002 \cdot T) \cdot (S - 35)]) \quad (4)$$

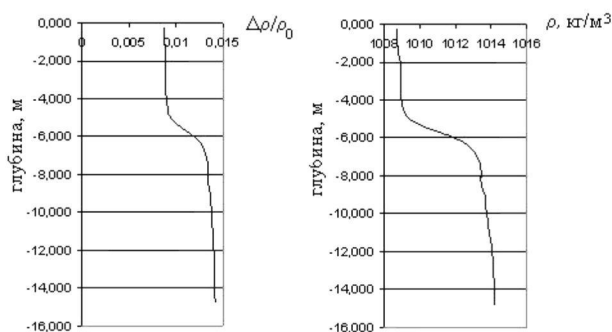


Рис. 5. Плотность, приращение плотности, формула (4)

Приращение  $\Delta\rho/\rho_0$  имеет значение 0,0054.

Далее значения для плотности и приращения плотности изображены на одном графике (рис. 6).

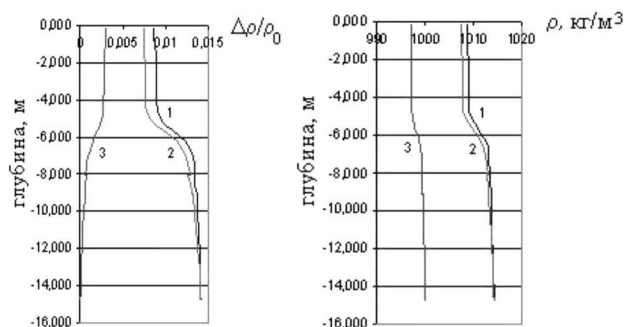


Рис. 6. Плотность, приращение плотности (1 — формула Мамаева, 2 — формула Бусси́неска, 3 — уравнение состояния равномерно соленой воды)

В некотором смысле «наиболее стратифицированное озеро» получается при использовании формулы Бусси́неска (рис. 7).



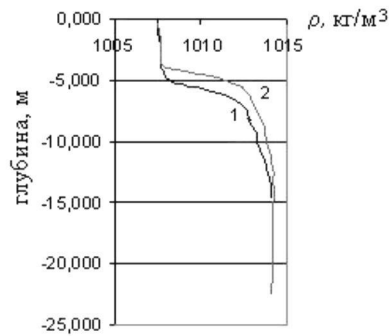


Рис. 7. Значение плотности, посчитанное по формуле Буссинеска для двух различных точек озера: 1 — плотность при измерениях 2 июля 2010 г. в точке глубины 14 м; 2 — плотность при измерениях 4 июля 2010 г. в центральной точке озера

Видно, что в течение двух дней в озере образовалась и вертикальная и горизонтальная плотностная неоднородность.

## 2. Анализ характеристик озера на основе измерения скоростей

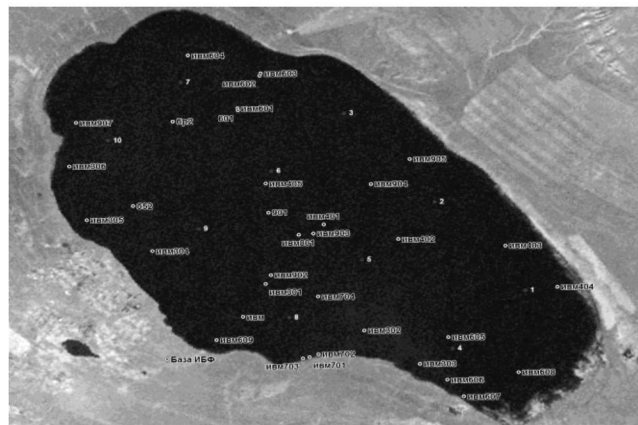


Рис. 8. Точки, в которых проводились измерения скорости течения

При выборе точек мы руководствовались тем соображением, что при том программном обеспечении, которое сейчас имеется, каждое измерение требует значительных затрат времени и составить одновременную картину поведения скорости течения в целом по озеру сложно, но реально отследить поведение жидкости в 3-х различных областях. Первая область – глубоководная часть с глубинами 15–25 м, средняя часть – с глубинами 7–14 м и, наконец, мелководная зона с глубинами до 7 м.

В сезоне измерений 2010 г. ветровая картина на озере характеризовалась резкой нестационарностью. Ветер менялся от нулевого значения до порывов в 10 м/с, при этом величина и направление ветра менялись от точки к точке в один день измерений.

Для повышения точности измерения два прибора ставились в одной точке так, чтобы одно измерение проходило лучами, направленными вниз, а второе — вверх. Но приборы находились в нескольких метрах друг от друга по горизонтали.

Первая серия наблюдений – наблюдения в точках различной глубины [1].

### *Глубоководная зона:*

02.07.2010; точка ИВМ, ветер с-в 2–2,6 м/с, глубина 16 м, 1200 kHz смотрел вниз, 600 kHz на дне смотрел вверх

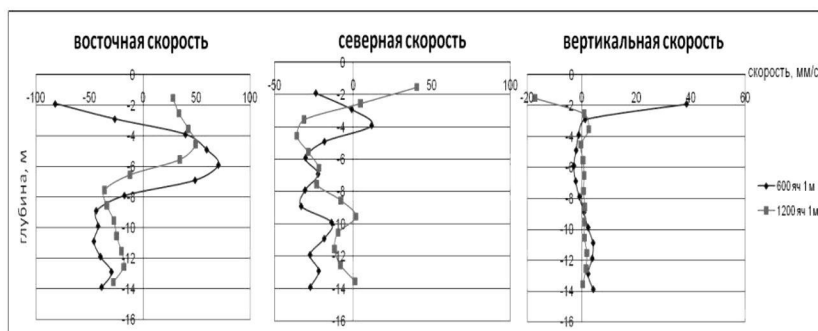


Рис. 9. Глубина 16 м

03.07.2010; точка ивм301, широта 54.29.855', долгота 90.11.828', ветер з - с-з 3 м/с с порывами до 5,5 м/с, глубина 22 м, смотрел вниз, 600 kHz на дне смотрел вверх

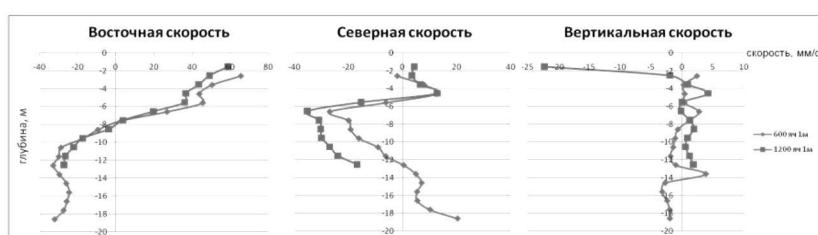


Рис. 10. Глубина 22 м

**Средняя зона:**

03.07.2010; точка ивм 302, ветер з от 0 до 1,5 м/с, через 5 мин от 2 м/с до 4 м/с, глубина 13 м, 1200 kHz смотрел вниз, 600 kHz на дне смотрел вверх

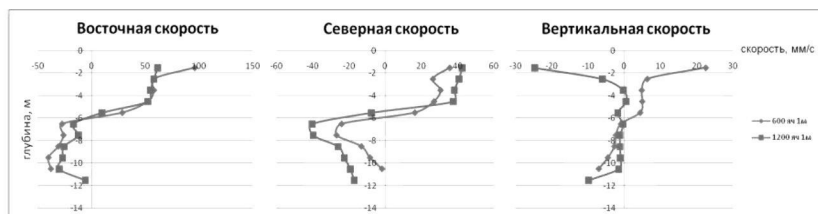


Рис. 11. Глубина 13 м

04.07.2010; точка ивм 402 (широта 54.30.318', долгота 90.13.680'), ветер ю-з 2 м/с, с порывами до 3,7 м/с, потом стих, глубина 13 м, 1200 смотрел вниз, 600 на дне смотрел вверх

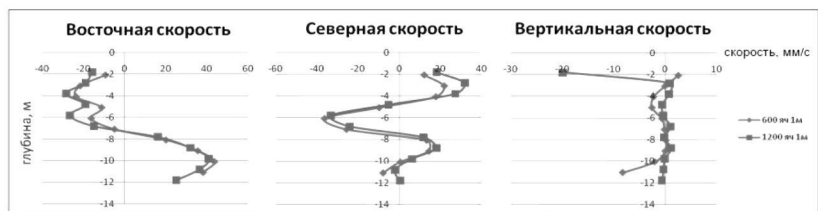


Рис. 12. Глубина 13 м

**Мелководная зона:**

6.07.2010; точка ивм606, ветер в 3-3,5 м/с с порывами до 5 м/с, глубина 5,7 м, 1200 смотрел вниз

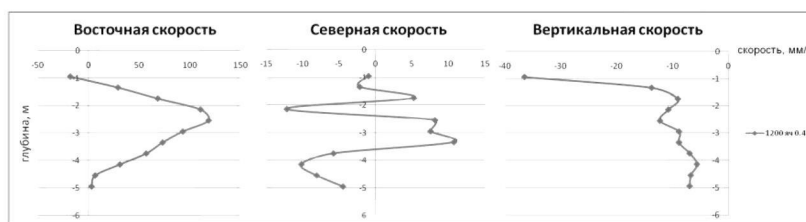


Рис. 13. Глубина 5,7 м

04.07.2010; точка ивм403 (широта 54.30.295', долгота 90.15.195'), ветер с-з 3,3 м/с с порывами до 4,5 м/с, глубина 5,5 м, 1200 смотрел вниз, 600 на дне смотрел вверх

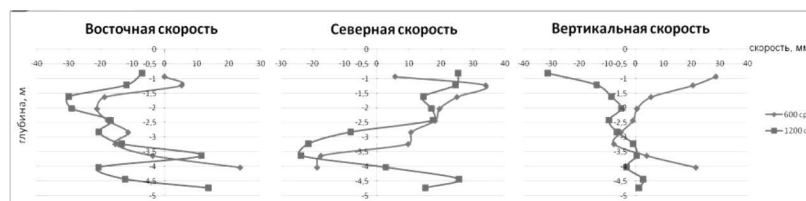


Рис. 14. Глубина 5,5 м

Вторая серия наблюдений относится к длительным наблюдениям в одной точке при быстро меняющемся ветре.

4.07.2010; точка ивм 405, ветер з - с-з 6 м/с с порывами до 10 м/с, глубина 24 м, 1200 кНз смотрел вниз, усреднение по времени 5 мин



Рис. 15. Глубина 24 м, усреднение 5 мин

В сезоне 2011 года проводились длительные измерения в центре озера.

Обработка данных с ADCP 600, полученных с 28 июня 2011 г. по 29 июня 2011 г. в результате длительного измерения (13 часов, прибор стоял на дне, центральная часть озера) показывает, что на глубине 12 м сильно колеблется вертикальная скорость, что является показательным при наличии в озере внутренних волн.

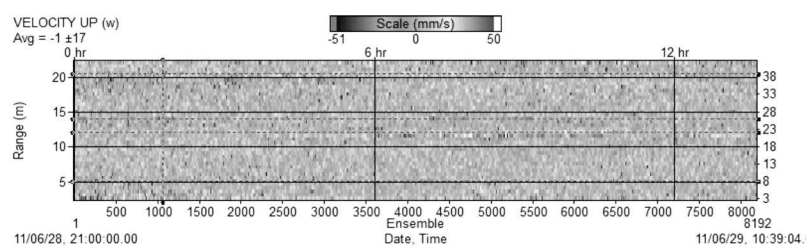


Рис. 16. Данные с ADCP 600, длительное измерение (13 часов с 28.06.11 по 29.06.11)

### 3. Оценка возможности появления внутренних волн в летний период в озере Шира

После обработки результатов длительных измерений 2011 г. усредненные данные (среднее за 1 минуту и 2 минуты) показаны на рис. 17, 18.

На обоих графиках видны колебания вертикальной скорости с периодом приблизительно 2 часа.



Рис. 17. Усредненные данные



Рис. 18. Усредненные данные

Сравним эти результаты с колебаниями температуры (данные А.П. Толмеева 2011 г.), измерения проводились на расстоянии 100 м от места, где измерялась скорость. По измерениям температура на этой глубине колебалась в рассматриваемый период от 1,0 °С до 1,5 °С, период колебаний температуры приблизительно 6 часов.

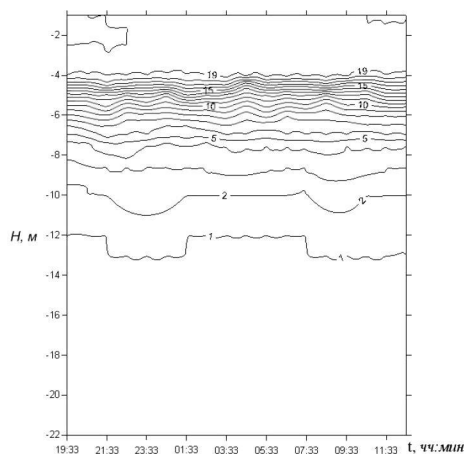


Рис. 19. Изотермы для периода длительных наблюдений с 28.06.2011 по 29.06.2011 (результат А.П. Толмеева)

Этот рисунок полностью идентичен рис. 20 [2], на котором схематически изображено изменение температуры при наличии в озере внутренних волн.

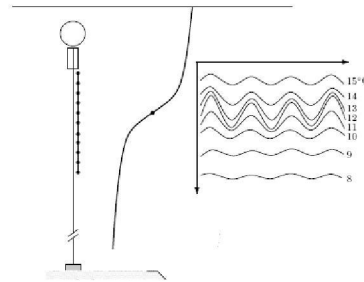


Рис. 20. Типичные изотермы для периода длительных наблюдений при наличии внутренних волн. Рисунок взят из книги [2].

Исходя из результатов книги [2], можно сделать вывод, что внутренние волны могут появляться в определенной области.

Проведем теоретический анализ возможных внутренних волн в озере Шира.

В соответствии с [2] область, в которой возможно появление внутренних волн, отражена на рис. 21.

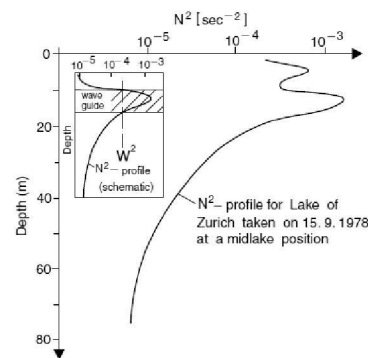


Рис. 21. Профиль квадрата частоты Брунта-Вайсяля, озеро Zurich [6]

Для озера Шира с использованием уравнений состояния (2) – (4) имеем следующие значения для величины  $N^2$

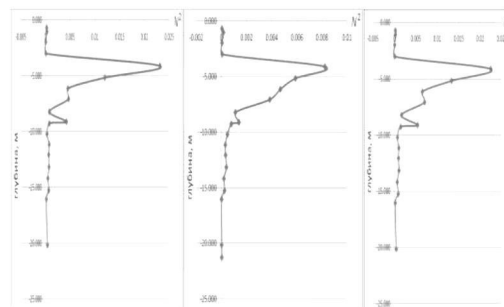


Рис. 22. 1-величина  $N^2$  в соответствии с формулой для равномерно соленой воды (2); 2- величина  $N^2$  в соответствии с формулой Буссинеска (3); 3- величина  $N^2$  в соответствии с формулой Мамаева (4)

Тогда можно ожидать появления внутренних волн в диапазоне 5-15 м, что не противоречит натурным данным.

### Заключение

Интерес к функционированию озерных систем велик во всем мире, при этом делаются попытки классифицировать озера по различным признакам. Например, в соответствии с тем, как они перемешиваются и т.д. Не на все эти вопросы есть сейчас ответ, но некоторые выводы сделать можно.

В рамках натурального изучения скоростного режима озера получены следующие результаты.

1. При измерении скорости в глубоководной зоне одна или обе скорости при входе в металимнион меняют свой знак.

2. Результаты натуральных измерений показывают, что в прибрежной зоне горизонтальная скорость может сохранять свой знак от поверхности до глубины, при этом возрастает порядок вертикальной скорости.

Как показывает анализ натуральных данных и аналитических решений, один из возможных и наиболее вероятных сценариев течения в озере Шира в летний период – это течение при наличии внутренних волн, тогда горизонтальные скорости имеют место вплоть до дна. Т.е. согласно международной терминологии озеро Шира в летний период – озеро-осциллятор [2].

#### Литература

1. Компаниец Л.А., Якубайлик Т.В., Питальская О.С. Опыт использования современных измерительных приборов для определения гидродинамических характеристик водоема // Прикладные технологии гидроакустики и гидрофизики: тр. X всерос. конф. СПб.: Наука, 2010. С. 310–311.
2. Waves and oscillations in the ocean and in lakes. In Continuum mechanics in environmental science and geophysics / ed. by K. Hutter – Springer-Verlag Berlin, New York, 1993. – 434 p.

*Компаниец Лидия Алексеевна*, канд. физ.-мат. наук, доц., старший научный сотрудник ИВМ СО РАН (г. Красноярск), тел. (391-2) 498811, [kla@icm.krasn.ru](mailto:kla@icm.krasn.ru)

*Якубайлик Татьяна Валерьевна*, младший научный сотрудник ИВМ СО РАН (г. Красноярск)

*Володько Ольга Станиславовна*, зав. лаб. Балтийского федерального университета им. И. Канта.

*Компаниец Lidiya Alekseevna*, candidate of physical and mathematical sciences, senior researcher ICM SB RAS (Krasnoyarsk), тел. (391-2) 498811, e-mail [rla@icm.krasn.ru](mailto:rla@icm.krasn.ru).

*Yakubaylik Tatiana Valerievna*, young researcher ICM SB RAS (Krasnoyarsk).

*Volodko Olga Stanislavovna*, Head of the Laboratory of the Immanuel Kant Baltic federal university.

**ОПЫТ РАЗРАБОТКИ И РЕЗУЛЬТАТЫ ПРИМЕНЕНИЯ  
ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА  
ДЛЯ МОДЕЛИРОВАНИЯ ГИДРОДИНАМИКИ И КАЧЕСТВА ВОДЫ РЕК  
И ВОДОЕМОВ БАСЕЙНА ОБИ.  
ЧАСТЬ 1 - УЧАСТОК РЕКИ СЕВЕРНАЯ СОСЬВА**

Дано описание состава, возможностей и ограничений применения информационно-вычислительного комплекса для численного моделирования гидродинамики и качества воды в предположении допустимости поперечного осреднения искомым переменных. Изложены конкретные результаты построения и верификации информационно-вычислительного комплекса для описания переменных гидродинамики участка реки Северная Сосьва, бассейн которой станет основным объектом техногенной нагрузки при реализации мега-проекта «Урал Промышленный – Урал Полярный».

**Ключевые слова:** река Северная Сосьва, база данных, модель гидродинамики и качества воды, верификация модели.

*P. Yu. Pushistov, V.N. Danchev, R.D. Romanenko*

**EXPERIENCE AND RESULTS OF THE DEVELOPMENT OF INFORMATION  
AND COMPUTATIONAL COMPLEX FOR SIMULATION OF HYDRODYNAMICS  
AND WATER QUALITY OF RIVERS AND LAKES OB BASIN.  
PART 1 - SECTION OF THE RIVER SEVERNAYA SOSVA**

The paper contains a description of the composition, capabilities and limitations of information and computing system for numerical simulation of fluid flow and water quality, assuming the lateral homogeneity of the unknown variables. Presented results of construction and verification of information and computing system to describe the hydrodynamic variables for the section river Severnaya Sosva, a basin which will be the main object of man-made loading in the implementation of mega-project "Ural Industrial - Ural Polar".

**Keywords:** river Severnaya Sosva, database, model of hydrodynamics and water quality, verification of the model.

**Введение**

Разработка информационно-вычислительных комплексов (ИВК) для моделирования гидродинамики и качества воды конкретных водных объектов (речные системы, озера, водохранилища и эстуарии) в составе инновационных систем поддержки принятия решений при интегрированном управлении речными бассейнами (СППР ИУРБ) требует последовательного применения методов системного анализа [1-3]. В ходе практической реализации основных этапов системного анализа (реальная водная экосистема → экспериментальное и теоретическое изучение системы → феноменологическая модель → математическая модель → имитационная компьютерная модель → верификация, анализ чувствительности и прогностических качеств численной модели) [4], при планировании и осуществлении конкретных проектов СППР ИУРБ особенно важной оказывается задача синтеза и целевого интегрирования программ сбора экосистемных данных, включая проведение полевых исследований, и программ разработки/адаптации информационно-вычислительного инструментария для создания адекватных численных моделей гидродинамики и качества воды [2, 5].

В настоящей публикации, состоящей из двух частей, излагается накопленный авторами опыт разработки и применения ИВК, базовым прогностическим модулем которого является модель гидродинамики и качества воды SE-QUAL-W2 [6-8], к двум конкретным водным объектам: участку реки Северная Сосьва (часть 1) и Телецкому озеру с главным притоком – рекой Чулышман (часть 2). Разработка модели SE-QUAL-W2 (далее W2) была инициирована Инженерным корпусом армии США в 1974 году [8]. Далее модель W2 непрерывно развивалась и успешно применялась на практике для решения широкого класса задач управления использованием и охраной водных ресурсов группой исследования качества воды Портландского государственного университета совместно с Инженерным корпусом армии США [6] и группой моделирования поверхностных вод компании ERM (управление ресурсами окружающей среды) [7]. На сегодня, согласно обобщенным данным [6], W2 использовалась для моделирования большого числа водных объектов (436 рек, 297 озер, 319 водохранилищ и 82 эстуария) в 116 странах.

Цель данной статьи состоит в том, чтобы дать краткую информацию о разработке и результатах применения ИВК в составе: база данных (БД), система моделирования бассейна (WMS) и модель W2 для расчета гидродинамики и качества воды участка реки Северная Сосьва (крупный левый приток реки Оби с длиной русла 720 км и площадью бассейна 89 тыс. км<sup>2</sup>, расположенный на территории Ханты-Мансийского автономного округа-Югры). Выбор указанного водного объекта продиктован следующими причинами:

1. Река Северная Сосьва образует уникальную естественно-природную экосистему с эндемиками в составе биоты (в их числе тугун – знаменитая сосьвинская селёдка), с нерестилищами особо ценных видов рыб, в том числе занесенных в Красную книгу РФ и ХМАО-Югры. В бассейне реки расположены природные территории, отнесенные к категории особо охраняемых, а также родовые угодья представителей коренных малочисленных народов Севера [9];

2. Бассейн верхнего течения реки в ближайшие годы станет зоной активного транспортно-энергетического и промышленного освоения (мегапроект «Урал Промышленный – Урал Полярный») [10], что делает важной задачу разработки проекта СППР ИУРБ Северной Сосьвы [11];

3. Югорским НИИ информационных технологий и Югорским государственным университетом в период 2004-2006 гг. накоплен положительный опыт применения модели W2, версия 3.1, для воспроизведения гидрологического, термического и гидрохимического режимов участка среднего течения реки Северная Сосьва от гидропоста Сосьва до гидропоста Сартынья с использованием данных наблюдений сети Росгидромета, а также данных батиметрии русла с лоцманских карт [12].

Принципиальным недостатком набора данных, использованного в то время при применении модели W2 для указанного участка реки Северная Сосьва [12], являлось крайне ограниченное количество данных о глубинах с лоцманских карт для формирования цифровой модели русла. Работа по подготовке файла батиметрии для W2 проводилась вручную и занимала огромное количество времени. Полностью отсутствовали данные гидрологических и гидрохимических наблюдений на участке реки между г/п Сосьва и г/п Сартынья, необходимые для полноценной калибровки и верификации численной модели.

Осенью 2007 года были проведены комплексные экспедиционные исследования на реке Северная Сосьва, программно ориентированные на использование полученных данных для разработки, калибровки и верификации новой версии модели W2. Результаты полевых работ, а также параметры и результаты верификации модифицированной версии модели W2 для участка реки от г/п Сосьва до г/п Сартынья (далее, модель W2-2009) достаточно полно описаны в [5]. Однако в модели W2-2009, как и в первой версии модели W2-2006 [12], файл батиметрии формировался преимущественно вручную, что и привело к необходимости разработки ИВК, компоненты которого представлены в следующем разделе статьи.

### 1. Состав информационно-вычислительного комплекса «Северная Сосьва»

Информационно-вычислительный комплекс, разработанный для моделирования гидродинамики и качества воды участка среднего течения реки Северная Сосьва (ИВК «Северная Сосьва»), включает в себя три модуля: 1 – WMS, 2 – W2 и 3 – БД для обеспечения работы модулей 1 и 2.

Прогностическим ядром ИВК «Северная Сосьва» является версия 3.2 численной продольно-вертикальной, поперечно осредненной модели W2 [8]. Гидродинамический блок модели W2 позволяет проводить с высоким пространственно-временным разрешением расчеты уровня поверхности воды, полей продольной и вертикальной составляющих скорости течения, температуры и турбулентности. Из-за использования предположения о поперечной однородности, наилучшими объектами применения этой модели являются относительно вытянутые и узкие водные объекты, характеризующиеся существенными продольными и вертикальными градиентами искомых переменных. Модель позволяет учитывать реальную внутрисуточную динамику метеорологических параметров (скорость и направление ветра, температура и влажность воздуха, балл облачности, солнечное излучение), а также параметры ветрового и светового затенения, связанного с топографией и растительностью береговой зоны. Модель рассчитывает характеристики турбулентности (всего версия 3.2 включает пять подмоделей расчета коэффициентов вертикального турбулентного обмена) и гидрооптические переменные водного объекта, а также описывает динамику ледового покрова. Блок качества воды версии 3.2, модульный по своей структуре, позволяет проводить расчеты 23 основных показателей качества воды (включая процессы эвтрофикации, представленные цепочкой: биогены → водоросли → растворенный кислород, органическое вещество, донные отложения, взвеси и т.п.) и более 60 производных от переменных основного состояния. В модели



также учитываются: боковая приточность (скорость, температура и характеристики качества воды); техногенные нагрузки от точечных и рассредоточенных источников (сбросов) загрязняющих веществ; процессы на границе раздела вода – седименты (наносы) и фильтрационный приток/отток подземных вод. Детальное описание возможностей и ограничений применимости указанной версии модели W2, заинтересованный читатель может найти в [8], а также в разделе «отчеты» [6]. Руководство [8], а также технические отчеты по применению модели W2 для рек Спокейн, Грин и Вилламетта, переведены первым автором данной статьи на русский язык.

Модель W2 имеет открытый исходный фортран-код и файловую структуру [8]. В связи с этим применение численной модели W2 для конкретного водного объекта требует квалифицированного заполнения трех основных файлов: контрольный, метеорологический и файл батиметрии. Контрольный файл содержит все средства управления, которые «указывают» численной модели W2, что выбрать в качестве начальных и граничных условий, какие файлы служат для просмотра результатов, на какой временной период проводится моделирование, какие численные схемы, подмодели и коэффициенты будут использоваться. Метеорологический файл прост по форме и содержит данные о значениях метеорологических величин, полученных с ближайших к водному объекту пунктов наблюдений. Файл батиметрии, на основе которого генерируется сетка численной модели, требует введения следующих видов информации о геометрии расчетной области: 1) длина сегментов, на которые разбиты водотоки и водные объекты; 2) данные о пространственной ориентации сегментов; 3) толщина слоев, которыми аппроксимируются поперечные сечения русел и пойм водотоков; 4) ширина слоев, упомянутых в 3.

В разработанном нами ИВК файл батиметрии модели W2 генерируется с помощью системы моделирования бассейна - WMS [13], созданной исследовательской лабораторией экологического моделирования университета Брайхема Юнга, совместно с Инженерным корпусом армии США. Одно из достоинств WMS заключается в ее способности проводить автоматизированную обработку разнообразной картографической информации, необходимой для создания файлов батиметрии модели W2, включая результаты специализированных полевых работ. Преимущество этого подхода заключается в том, что сгенерированный файл батиметрии для модели W2 позволяет более полно и достоверно учесть реальные особенности морфометрии русла (ложа) водного объекта.

База данных, необходимых для формирования основных файлов (контрольного, метеорологического, батиметрии) и дополнительных файлов для задания соответствующих начальных и граничных условий, условно разделена на две части. Первая часть БД содержит гидрометеорологическую и гидрохимическую информацию с пунктов наблюдений Росгидромета (метеорологические данные с м/с Сосьва, гидрологические данные: уровень, расход и температура воды с г/п Сосьва и г/п Сартынья, гидрохимические данные с г/п Сосьва за 2003-2004 гг.), а также картографическую информацию (лоцманская карта участка реки, топографическая карта соответствующей береговой зоны масштабов 1:25000). Вторая часть БД сформирована на основе данных полевых работ, проведенных в сентябре 2007 года на участке среднего течения реки Северная Сосьва [5]. Подчеркнем, что основная цель этих работ заключалась в получении гидрологических (уровни поверхности воды -  $H$ , скорости течения -  $U$  и температура воды -  $T$ ) и гидрографических (глубины -  $Z$ ) данных высокого пространственного разрешения (всего 31 промерный створ на участке реки длиной 70 км), необходимых для уточнения входных параметров модели W2, а также для верификации результатов расчетов.

## **2. Технология и результаты автоматизированной генерации модели батиметрии с использованием картографических материалов и данных полевых работ**

При разработке версии модели W2 для участка реки Северная Сосьва от г/п Сосьва до г/п Сартынья в составе ИВК «Северная Сосьва» (далее модель W2-2011) с усвоением данных второй части БД, первоочередной являлась задача автоматизированного построения файла батиметрии, необходимого для генерации расчетной сетки в W2-2011. Ключевым моментом в построении батиметрии с использованием WMS состоит в том, что необходимо получить нерегулярную триангуляционную сеть (TIN) отметок высот над уровнем моря, описывающую русловую и пойменную части указанного участка реки, на основе которой будет сгенерирована цифровая модель рельефа. Для этого были использованы следующие наборы данных:

- 1) номенклатурные листы топографической карты участка береговой зоны реки масштаба 1:25000;
- 2) лоцманские карты реки Северная Сосьва на исследуемом участке;

3) данные эхолотирования и поперечные разрезы русла, полученные в результате проведения полевых работ 2007 г.

На топографических картах в качестве исходных данных принимались: горизонтали высот, высотные отметки, урезы воды, линейные элементы гидросети и контуры водотоков, оцифровка которых осуществлялась с применением программ Easy Trace 7.99 и ArcGis 9.3 в проекции Гаусса—Крюгера (Пулково-42), зона 11. Лоцманские карты были привязаны к указанной проекции и оцифрованы (урезы воды, отметки глубин и изобаты) в ArcGis 9.3. Из набора данных полевых работ были отобраны данные эхолотирования и поперечные разрезы русла, имеющие координатную привязку. Данные об отметках высот и глубин, содержащиеся в перечисленных выше источниках, с помощью ArcGIS 9.3 были представлены в виде TIN, описывающей русловую и пойменную части указанного участка реки.

На основе TIN создается цифровая модель рельефа в WMS, которая представляет собой описание исследуемой территории, включающее в себя такие геометрические характеристики, как морфометрические особенности русла и берега, границу моделируемой территории, маршрутизацию потока и т.п. После создания цифровой модели рельефа производится декомпозиция водного объекта на водотоки и сегменты. Далее необходимые для расчетов данные конвертируются WMS в элементы и узлы сетки модели W2-2011. Затем для каждого сегмента генерируется объемная кривая, которая определяет отношение глубины к объему водотока в каждом сегменте, и с ее помощью задаются толщины слоев и вычисляются ширины ячеек. Если в модель батиметрии при необходимости вносятся какие-либо изменения, то сама сетка и ее атрибуты легко и быстро обновляются.

Сетка модели W2-2011, полученная с помощью ИВК «Северная Сосьва» (см. рис. 1), имеет следующие параметры: общее число сегментов – 450, в том числе, 357 сегментов основного русла и 93 сегмента, приходящиеся на участки реки, отделенные от основного русла островами. Всего в сетку модели W2-2011 включено семь протоков. Длины сегментов модели W2-2011 изменяются в диапазоне от 47.6 м до 271.5 м при средних значениях около 200 м, число вертикальных слоев равно 104, толщина слоев – 0,2 м. Для удобства сравнения различий в сеточных представлениях моделей W2-2011 и W2-2009 на рис. 2в приведена схема расчетной сетки модели W2-2009 с одинаковыми длинами сегментов, равными 400 м (число слоев по вертикали – 104, толщина слоев – 0,2 м).

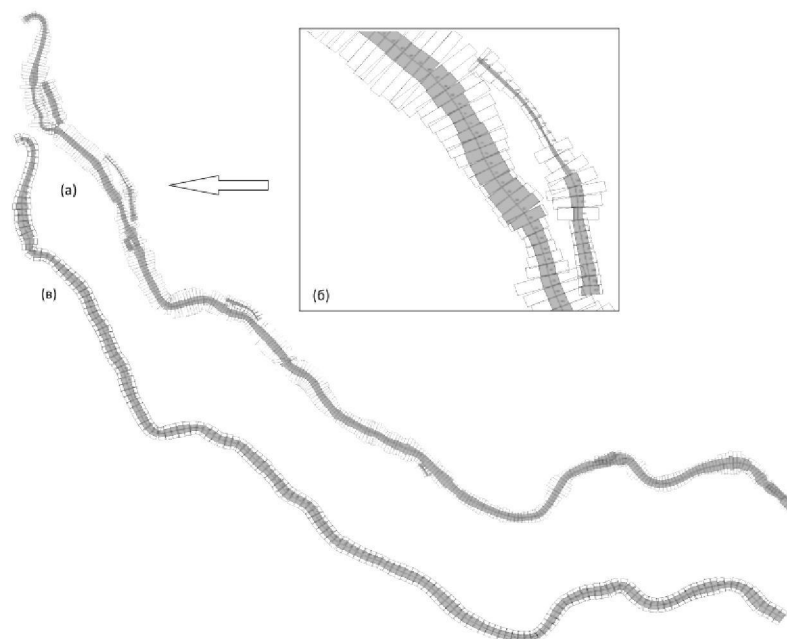


Рис 1. Схема расчетной сетки в плане участка реки Северная Сосьва от г/п Сосьва до г/п Сартынья: (а) – модель W2-2011; (б) – остров и проток Яныг-Пальятумп; (в) - модель W2-2009.

### 3. Результаты верификации ИВК «Северная Сосьва»

Верификация разработанного ИВК «Северная Сосьва» проводилась на основе сравнения результатов численных экспериментов с блоком гидродинамики модели W2-2011 и данных полевых работ 2007 года (вторая часть БД), а также посредством сопоставительного анализа результатов

расчета по моделям W2-2009 и W2-2011. Период моделирования совпадал с периодом выполнения полевых работ на участке реки от г/п Сосьва до г/п Сартынья – 20-21 сентября 2007 года. Краевые условия для Н и Т на входном и выходном створах задавались по данным измерений на г/п Сосьва и г/п Сартынья. Начальные условия для Н рассчитывались с помощью линейной интерполяции значений Н на входном и выходном створах. Краевые условия для Т на поверхности воды рассчитывались с помощью подмодели ЕТ (приближение равновесной температуры, см. [8]), коэффициенты Маннинга в параметризации придонного трения были заданы равными 0,03 в обеих моделях. Одинаковыми были и модели, используемые для описания поля коэффициентов вертикального турбулентного обмена (модель Коли и Бучака [8]).

Общее представление о качестве моделирования уровня поверхности воды (Н) вдоль участка реки дает рис. 2. В табл. 1 приведены значения абсолютных ( $\Delta_H = |H_H - H_M|$ , где  $H_H$  – данные наблюдений, полученные с помощью аппаратуры Trimble 5700 и  $H_M$  – результаты моделирования с помощью W2-2011) и относительных ( $\delta_H = \Delta / Z \times 100\%$ ) ошибок в определении уровней поверхности воды за 20-21 сентября 2007 г.

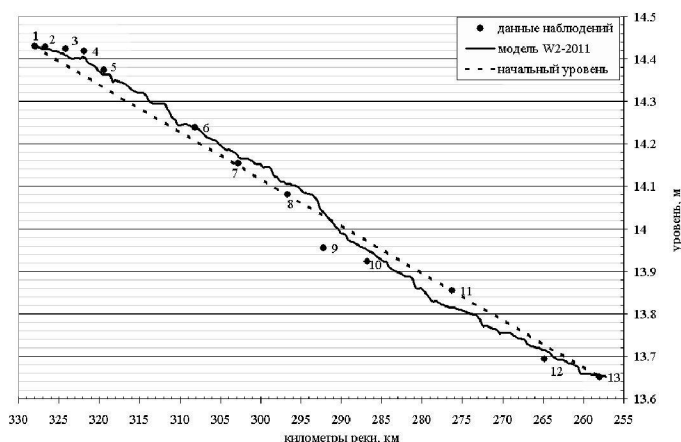


Рис 2. Результаты моделирования продольного профиля уровня воды на участке реки Северная Сосьва (сплошная кривая) и данные наблюдений за 20-21.09.2007 г. (точки).

Таблица 1

Погрешности результатов моделирования уровня воды за 20-21.09.2007 г. ( $\mu$  – средние значения)

№	Z, м	$\Delta_H$ , м	$\delta_H$ , %
1	6,15	0,002	0,03
2	12,15	0,004	0,03
3	9,15	0,001	0,01
4	7,55	0,015	0,20
5	7,95	0,011	0,14
6	4,35	0,001	0,01
7	5,35	0,013	0,25
8	4,55	0,031	0,67
9	4,35	0,416	9,55
10	5,95	0,031	0,52
11	9,15	0,043	0,47
12	7,35	0,020	0,27
13	6,55	0,005	0,08
$\mu$	<b>6,96</b>	<b>0,046</b>	<b>0,94</b>

Из рис. 2 и табл. 1 можно сделать вывод о достаточно высокой точности воспроизведения Н в верхней половине участка реки (точки 2-8) и о понижении точности в нижней половине участка (точки 9-11).

Рис. 3 дает конкретное представление об изменчивости рассчитанных по моделям W2-2011 (рис. 3,а) и W2-2009 (рис. 3,б) значений U в поверхностном слое воды вдоль участка реки (сплошная кривая) и поперечно-осредненных в каждом створе данных измерений скоростей течений, по-

лученных с помощью акустического измерителя скорости течения и глубин «Mini» ADP (точки) за 20-21 сентября 2007 г.

Табл. 2 содержит значения абсолютных ( $\Delta_u = |U_H - U_M|$ , где  $U_H$  – измеренные, а  $U_M$  – рассчитанные значения  $U$ ) и относительных ( $\delta_u = \Delta_u / U_H \times 100\%$ ) ошибок, полученных с помощью моделей W2-2011 и W2-2009, соответственно.

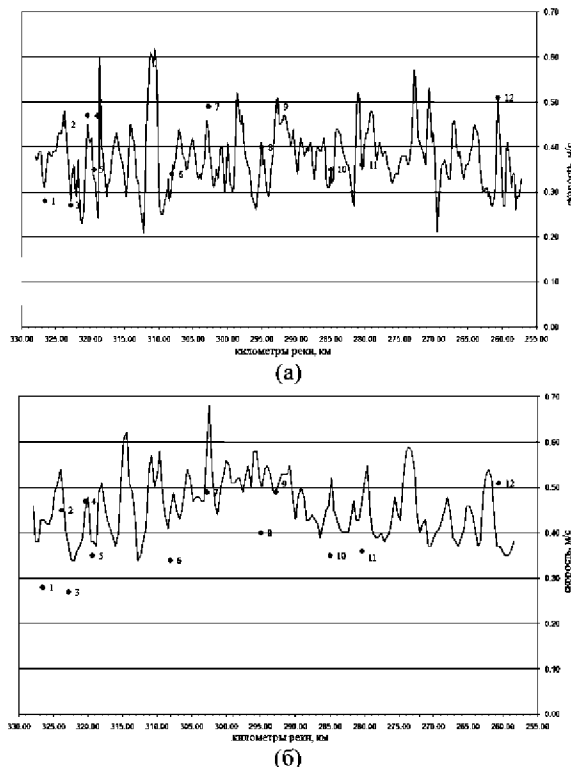


Рис. 3. Результаты моделирования продольной скорости течения в поверхностном слое воды за 20-21.09.2007 г.: а – W2-2011, б – W2-2009.

Таблица 2  
Погрешности результатов моделирования продольной скорости течения за 20-21.09.2007 г.

№	модель W2-2011		модель W2-2009	
	$\Delta_u, \text{ м/с}$	$\delta_u, \%$	$\Delta_u, \text{ м/с}$	$\delta_u, \%$
1	0.040	14.29	0,150	53,57
2	0.020	4.44	0,065	14,44
3	0.000	0.00	0,100	37,04
4	0.020	4.26	0,020	4,26
5	0.020	5.71	0,030	8,57
6	0.010	2.94	0,120	35,29
7	0.040	8.16	0,100	20,41
8	0.010	2.50	0,110	27,50
9	0.020	4.08	0,000	0,00
10	0.000	0.00	0,140	40,00
11	0.010	2.78	0,110	30,56
12	0.000	0.00	0,140	27,45
$\mu$	<b>0.016</b>	<b>4.10</b>	<b>0,090</b>	<b>24,92</b>

Сопоставление данных табл. 2 свидетельствует о заметном преимуществе модели W2-2011 по сравнению с моделью W2-2009 в точности описания наблюдаемых значений  $U$  в поверхностном слое воды.

К сожалению, ситуация с данными гидрохимических и гидробиологических наблюдений на изучаемом участке среднего течения реки Северная Сосьва не изменилась к лучшему с момента публикации [12]. Есть данные стандартных гидрохимических наблюдений (5-6 раз в год) на г/п

Сосьва, однако аналогичные наблюдения на г/п Сартынья не проводятся, поэтому верификация блока качества воды модели W2-2011 по существу оказалась невозможной.

### **Заключение**

1. На основе интегрирования системы WMS и модели CE-QUAL-W2 разработан единый информационно-вычислительный комплекс для моделирования гидродинамики и качества воды участка среднего течения реки Северная Сосьва, бассейн которой станет основным объектом технологической нагрузки при реализации мегапроекта «Урал Промышленный – Урал Полярный».

2. Информационно-вычислительный комплекс «Северная Сосьва» позволяет автоматизировать разработку модели батиметрии водного объекта, что существенно сокращает долю ручного труда при подготовке файла батиметрии. Одновременно, с помощью этого комплекса удастся увеличить точность воспроизведения переменных гидродинамики по сравнению с ранее разработанными моделями.

3. Описанный в статье информационно-вычислительный комплекс, по мнению авторов, может быть адаптирован и применен для других крупных рек России при решении прикладных задач использования, охраны и восстановления водных ресурсов, в том числе, при разработке схем комплексного использования и охраны водных объектов (СКИОВО) и нормативов допустимого воздействия на водные объекты (НДВ).

### **Благодарности**

Авторы выражают благодарность чл.-кор. РАН В.Н. Лыковому и проф. В.А. Земцову за ценные советы и рекомендации по проблемам системного анализа водных объектов. Авторы признательны проф. С.А. Веллсу и доктору Бергеру К. за поддержку нашей работы и полезные консультации по вопросам применения базовых программных средств.

### **Литература**

1. Интегрированное управление водными ресурсами Санкт-Петербурга и Ленинградской области / А.Ф. Алимов и др. // Опыт создания системы поддержки принятия решений. – СПб.: Borey Print, 2001. - 419 с.
2. Daniel P. Loucks and Eelco van Beek with contributions from Jerry R. Stedinger, Jozef P.M. Dijkman, Monique T. Villars. Water Resources Systems Planning and Management An Introduction to Methods, Models and Applications. Studies and Reports in Hydrology. UNESCO PUBLISHING. – ISBN 92-3-103998-9 – © UNESCO 2005. - 676 p.
3. Водные ресурсы и качество вод: состояние и проблемы управления / В.И. Данилов-Данильян и др. М.: РАСХН, 2010. - 415 с.
4. Страшкраба М., Гнаука А. Пресноводные экосистемы. Математическое моделирование. М.: Мир, 1989. - 376 с.
5. Пушистов П.Ю., Вторушин М.Н., Романенко Р.Д. Результаты применения методов системного анализа при исследовании детальной пространственно-временной структуры переменных гидродинамики и качества воды среднего течения крупной реки Северная Сосьва // Водные проблемы крупных речных бассейнов и пути их решения: сб. науч. тр. Барнаул: ООО «Агентство рекламных технологий», 2009. С. 333-345.
6. URL: <http://www.ce.pdx.edu/w2/>
7. URL: <http://www.erm-smg.com/ce-qual-w2.html>
8. Cole T.M., Wells S.A. “CE-QUAL-W2” A two-dimensional, laterally averaged, Hydrodynamic and Water Quality Model, Version 3.2, Instruction Report EL-2003-1, US Army Engineering and Research Development Center, Vicksburg, MS, 2003. - 634 p.
9. Характеристика экосистемы реки Северной Сосьвы / под ред. Л.Н. Добринского. Свердловск: УрО АН СССР, 1990. - 255 с.
10. Концепция комплексного промышленного освоения Приполярного Урала на основе опережающего развития транспортной и энергетической инфраструктуры. Раздел «Недропользование» (краткое содержание). Ханты-Мансийск: ОАО НПЦ «Мониторинг», 2006, 39 с.
11. Пушистов П.Ю., Дикунец В.А., Земцов В.А., Романенко Р.Д. Исследование осуществимости разработки системы поддержки принятия решений по сохранению уникальной экосистемы реки Северная Сосьва при крупномасштабном горно-промышленном и транспортно-энергетическом освоении Приполярного Урала / Фундаментальные проблемы воды и водных ресурсов: материалы Третьей всерос. конф. с междунар. участием. Барнаул, 24-28 августа 2010 г. Барнаул: Изд-во АРТ, 2010. С. 516–519.
12. Численное моделирование пространственно-временной структуры гидродинамики и характеристик качества воды реки Северная Сосьва / П.Ю. Пушистов и др. // Оптика атмосферы и океана. 2006. Том 19. № 11. С. 956-960.
13. URL: <http://www.aquaveo.com>

*Пушистов Петр Юрьевич*, доктор физико-математических наук, профессор, научный консультант, Институт водных и экологических проблем Сибирского отделения Российской академии наук, Россия, 656038, г. Барнаул, ул. Молодежная, 1; e-mail: pushtcmpr@mail.ru

*Данчев Василий Николаевич*, аспирант, Институт водных и экологических проблем Сибирского отделения Российской академии наук, Россия, 656038, г. Барнаул, ул. Молодежная, 1; e-mail: danchev\_vasilii@mail.ru

*Романенко Роман Дмитриевич*, ведущий инженер, Территориальный центр анализа и прогноза угроз безопасности жизнедеятельности ХМАО-ЮГРЫ, Россия, 628012, г. Ханты-Мансийск, ул. Гагарина, 153/а; e-mail: rromanenkoRD@yandex.ruinfo

*Pushistov Peter Yurievich*, doctor of physical and mathematical sciences, professor, scientific adviser, Institute for Water and Environmental Problems Siberian Branch of the RAS, Russia, 656038, Barnaul, ul. Molodyoznaya, 1.

*Danchev Vasily Nikolayevich*, postgraduate, Institute for Water and Environmental Problems Siberian Branch of the RAS, Russia, 656038, Barnaul, ul. Molodyoznaya, 1.

*Romanenko Roman Dmitrievich*, senior engineer, Regional center of analysis and forecasting of threats to life safety Khanty-Ugra, Russia, 628012, Khanty-Mansiysk, ul. Gagarina, 153/a.

## СОДЕРЖАНИЕ

### 1. Управляемые системы и методы оптимизации

<i>Дранко О.И., Яковенко Г.Н.</i> Оптимизация дивидендной политики.....	3
<i>Колмакова А.И.</i> Многометодная оптимизация управления в экономической модели выбора налоговой ставки.....	9
<i>Мальшиев А.В.</i> Алгоритм глобального поиска гарантированных решений квадратично-линейной двухуровневой задачи и его тестирование.....	17

### 2. Системный анализ и информационные технологии

<i>Ветров А.А., Фереферов Е.С., Хмельнов А.Е., Шигаров А.О.</i> Формирование хранилища данных для системы MDAttr.....	22
<i>Кононов Д.Д., Исаев С.В.</i> Модель безопасности веб-приложений на основе мандатного ролевого разграничения доступа.....	29
<i>Кравченко В.А.</i> Моделирование поиска решения с помощью функциональных грамматик.....	33
<i>Лукьянова Л.М., Федорченко Л.Н.</i> Средства формализации целей и проблем сложных систем производственной сферы.....	42
<i>Сахабетдинов И.У.</i> Задачи комплексного мониторинга в автоматизированных системах диспетчерского управления энергетическими объектами.....	49
<i>Фереферов Е.С., Новицкий Ю.А., Ружников Г.М., Хмельнов А.Е.</i> Технология интеграции геоинформационных функций в информационные системы.....	59
<i>Харинов М.В.</i> Устойчивая сегментация изображения.....	64
<i>Чернышов М.Ю., Абасов Н.В., Осипчук Е.Н.</i> Основы вычислительной технологии для анализа и реструктурирования текстов программ на принципах логико-смыслового анализа и синтеза.....	70
<i>Чернышов М.Ю., Нурминский В.Н., Абасов Н.В., Осипчук Е.Н.</i> Метод количественного анализа объектов на видеоизображениях, рассчитанный на использование в ГИС-технологиях и анализе микроскопических видеоизображений.....	76
<i>Шелехов В.И., Тумуров Э.Г.</i> Логика невзаимодействующих программ и реактивных систем.....	81

### 3. Функциональный анализ и дифференциальные уравнения

<i>Аверина Т.А.</i> Устойчивые численные методы решения стохастических дифференциальных уравнений в смысле Стратоновича.....	91
<i>Баргуев С.Г., Аюшеев Т.В., Миждон А.Д.</i> Об одном обобщении для решения начально-краевой задачи о колебаниях произвольного числа осцилляторов на стержне.....	95
<i>Кибирев В.В.</i> Об одной задаче наклонной производной с линейными коэффициентами... ..	101
<i>Николаев О.Ю.</i> Разрешимость обратной задачи для параболического уравнения высокого порядка с неизвестным коэффициентом поглощения.....	103
<i>Новиков А.Е., Шайдулов В.В.</i> Алгоритм интегрирования переменной конфигурации на основе явно-неявных схем четвертого порядка.....	111
<i>Русанов В.А., Антонова Л.В., Данеев А.В., Миронов А.С.</i> О формально разрешимых реализациях динамических систем с минимальной операторной нормой.....	117

### 4. Математическое моделирование

<i>Архипов С.В., Брагин А.Ф.</i> Нейросетевой подход в построении адаптивных сетей для нерегулярных плоских областей.....	134
<i>Болдсайхан Б., Дармаев Т.Г., Хабитуев Б.В., Хандаров Ф.В.</i> Моделирование пульсовой волны по хаотическим рядам.....	142
<i>Воротынцева А.В.</i> Некоторые характерные графы работ для представления сценариев расчета математических моделей .....	147

<i>Данчев В.Н., Пушистов П.Ю.</i> Опыт разработки и результаты применения информационно-вычислительного комплекса для моделирования гидродинамики и качества воды рек и водоемов бассейна Оби. Часть 2 – Телецкое озеро и участок реки Чулышман.....	154
<i>Итигилов Г.Б., Ширанов Д.Ш.</i> Метод инвариантных преобразований для определения поперечных компонент электромагнитного поля в гиротропных ограниченных областях .....	162
<i>Копманиец Л.А., Якубайлик Т.В., Володько О.С.</i> Анализ характеристик озера Шира на основе натурных наблюдений .....	167
<i>Пушистов П.Ю., Данчев В.Н., Романенко Р.Д.</i> Опыт разработки и результаты применения информационно-вычислительного комплекса для моделирования гидродинамики и качества воды рек и водоемов бассейна Оби. Часть 1 – участок реки Северная Сосьва.....	177



## CONTENTS

### 1. Control systems and optimization methods

<i>Dranko O.I., Yakovenko G.N.</i> The optimization of dividend policy.....	3
<i>Kolmakova A.I.</i> Multimethod optimizing of control in the economic model of the choice of tax rate.....	9
<i>Malyshev A.V.</i> Algorithm of global search for guaranteed solutions to quadratic-linear bilevel problem and its testing.....	17

### 2. System analysis and information technologies

<i>Vetrov A.A., Fereferov E.S., Hmelnov A.E., Shigarov A.O.</i> Forming data warehouse for system MDAttr.....	22
<i>Kononov D.D., Isaev S.V.</i> Mandatory role-based access control for web applications.....	29
<i>Kravchenko V.A.</i> Modelling of search of the decision by functional grammars.....	33
<i>Lukianova L.M., Fedorchenko L.N.</i> The means of goals and problems formalization in industrial complexes.....	42
<i>Sakhabetdinov I.U.</i> Challenges of integrated monitoring in automated systems dispatching energy facilities.....	49
<i>Fereferov E.S., Novitsky Yu.A., Rugnikov G.M., Hmelnov A.E.</i> Technology of integration of GIS functions to information systems.....	59
<i>Kharinov M.V.</i> Stable image segmentation.....	64
<i>Chernyshov M.Yu., Abasov N.V., Osipchuk E.N.</i> Foundations of the computational technology intended for analysis and restructuring software texts on the principles of logic-semantic analysis and synthesis.....	70
<i>Chernyshov M.Yu., Nurminsky V.N., Abasov N.V., Osipchuk E.N.</i> The method of quantitative analysis of objects on viseoimages, which is intended for applications in gis-technologies and in analysis of microscopic viseoimages.....	76
<i>Shelekhov V.I., Tumurov E.G.</i> Logic of non-interacting programs and reactive systems.....	81

### 3. Functional analysis and differential equations

<i>Averina T.A.</i> Numerical methods for solving stochastic differential equations in the Straatonovich sense.....	91
<i>Barguev S.G., Ausheev T.V., Mizhidon A.D.</i> A generalization for solving initial-boundary problem about fluctuations of arbitrary number oscillators on the elastic core.....	95
<i>Kibirev V.V.</i> On one problem of the oblique derivative with linear coefficients.....	101
<i>Nikolaev O.Yu.</i> Resolvability of a return task for the parabolic equation of a high order with unknown factor of absorption.....	103
<i>Novikov A.E., Shaidurov V.V.</i> An integration algorithm of variable configuration based on explicit-implicit schemes of 4th order of accuracy.....	111
<i>Rusanov V.A., Antonova L.V., Daneev A.V., Mironov A.S.</i> On the formal realizations of solvable dynamical systems with minimal operator norm.....	117

### 4. Mathematical modeling

<i>Arkhipov S.V., Bragin A.F.</i> The neural network approach in construction of adaptive grids for irregular flat areas.....	134
<i>Boldsaikhan B., Darmaev T.G., Khabituev B.V., Khandarov F.V.</i> Modeling of pulse wave using the chaotic series.....	142
<i>Vorotyntsev A.V.</i> Some typical workflow graphs for representation of the scenarios of calculating the mathematical models.....	147
<i>Danchev V.N., Pushistov P.Yu.</i> Experience and results of the development of information and computational complex for simulation of hydrodynamics and water quality of rivers and lakes Ob basin. Part 2 - Lake Teletskoye and section of the river Chulyshman.....	154

<i>Itgilov G.B., Shirapov D.Sh.</i> Method of invariant transformations for definition cross-section a component of an electromagnetic field in girotropic the limited areas.....	162
<i>Kompaniets L.A., Yakubaylik T.V., Volodko O.S.</i> Analysis of the characteristics of the Shira lake on the basis of field observations.....	167
<i>Pushistov P. Yu., Danchev V.N., Romanenko R.D.</i> Experience and results of the development of information and computational complex for simulation of hydrodynamics and water quality of rivers and lakes Ob basin. Part 1 - section of the river Severnaya Sosva.....	177

Научное издание

ВЕСТНИК  
БУРЯТСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Выпуск 9

МАТЕМАТИКА, ИНФОРМАТИКА

Свидетельство о регистрации  
ПИ №ФС77–36152 от 06 мая 2009 г.  
Федеральная служба по надзору в сфере связи,  
информационных технологий  
и массовых коммуникаций (Роскомнадзор)

Редактор  
*Е.П. Евдокимова*

Компьютерная верстка  
*Д.О. Трунин, И.-Х.Д. Хишектуева*

Свидетельство о государственной аккредитации  
№ 1289 от 23 декабря 2011 г.

Подписано в печать 16.06.12. Формат 60x84 1/8.  
Уч.-изд. л. 21,39. Усл.печ. л. 22,21. Тираж 1000. Заказ 209.

Издательство Бурятского госуниверситета  
670000, г. Улан-Удэ, ул. Смолина, 24а  
riobsu@gmail.com

## ВЕСТНИК БУРЯТСКОГО ГОСУНИВЕРСИТЕТА

Вестник БГУ включен в подписной каталог Роспечати за № 18534 и Перечень изданий Российской Федерации, где должны быть опубликованы основные научные результаты диссертаций на соискание ученых степеней доктора и кандидата наук.

На основании постановления заседания Ученого совета БГУ за № 10 от 28 мая 2009 г. в «Вестнике БГУ» в 2012 г. публикуются статьи по следующим направлениям:

**1. Педагогика (январь)**

гл. ред. Дагбаева Нина Жамсуевна – тел. 21-04-11; 44-23-95

эл. адрес: [vestnik\\_pedagog@bsu.ru](mailto:vestnik_pedagog@bsu.ru)

**2. Экономика. Право (февраль)**

гл. ред. Бадмаева Мария Валентиновна – тел. 21-37-44

эл. адрес: [vestnik\\_econom@bsu.ru](mailto:vestnik_econom@bsu.ru)

**3. Химия, физика (март)**

гл. ред. Хахинов Вячеслав Викторович – тел. 43-42-58

эл. адрес: [khakhinov@mail.ru](mailto:khakhinov@mail.ru)

**4. Биология, география (март)**

гл. ред. Доржиев Цыдып Заятуевич – тел. 21-03-48

эл. адрес: [vestnik\\_biolog@bsu.ru](mailto:vestnik_biolog@bsu.ru)

**5. Психология, социальная работа (апрель)**

гл. ред. Базарова Татьяна Содномовна – тел. 21-26-49

эл. адрес: [decspf@mail.ru](mailto:decspf@mail.ru)

**6. Философия, социология, политология, культурология (апрель)**

гл. ред. Осинский Иван Иосифович – тел. 21-05-62

эл. адрес: [intellige2007@rambler.ru](mailto:intellige2007@rambler.ru)

**7. История (май)**

гл. ред. Митупов Константин Батомункич – тел. 21-64-47

эл. адрес: [vestnik\\_history@bsu.ru](mailto:vestnik_history@bsu.ru)

**8. Востоковедение (май)**

гл. ред. Бураев Дмитрий Игнатьевич – тел. 44-25-22

эл. адрес: [railia@mail.ru](mailto:railia@mail.ru)

**9. Математика, информатика (июнь)**

гл. ред. Булдаев Александр Сергеевич – тел. 21-97-57

эл. адрес: [vestnik\\_matem@bsu.ru](mailto:vestnik_matem@bsu.ru)

**10. Филология (сентябрь)**

гл. ред. Имixelова Светлана Степановна – тел. 21-05-91

эл. адрес: [vestnik\\_phylolog@bsu.ru](mailto:vestnik_phylolog@bsu.ru)

**11. Романо-германская филология (сентябрь)**

гл. ред. Ковалева Лариса Петровна – тел. 21-17-98

эл. адрес: [klp@bsu.ru](mailto:klp@bsu.ru), [khida@mail.ru](mailto:khida@mail.ru)

**12. Медицина, фармация (октябрь)**

гл. ред. Хитрихеев Владимир Евгеньевич – тел. 44-82-55

эл. адрес: [vestnik\\_medicine@bsu.ru](mailto:vestnik_medicine@bsu.ru)

**13. Физкультура и спорт (октябрь)**

гл. ред. Гаськов Алексей Владимирович – тел. 21-69-89

эл. адрес: [gaskov@bsu.ru](mailto:gaskov@bsu.ru)

**14. Философия, социология, политология, культурология (ноябрь)**

гл. ред. Осинский Иван Иосифович – тел. 21-05-62

эл. адрес: [intellige2007@rambler.ru](mailto:intellige2007@rambler.ru)

**15. Теория и методика обучения (декабрь)**

гл. ред. Очиров Михаил Надмитович – тел. 21-97-57

эл. адрес: [vestnik\\_method@bsu.ru](mailto:vestnik_method@bsu.ru)

## Требования к оформлению статей, представляемых в «Вестник БГУ»

Отбор и редактирование публикуемых статей производятся редакционной коллегией из ведущих ученых и приглашенных специалистов.

В «Вестник БГУ» следует направлять статьи, отличающиеся высокой степенью научной новизны и значимостью. Каждая статья имеет УДК, а также письменный развернутый отзыв (рецензию) научного руководителя или научного консультанта, заверенный печатью.

Общие требования	Тексты представляются в электронном и печатном виде. Файл со статьей может быть на дискете или отправлен электронным письмом. На последней странице – подпись автора(ов) статьи. Название статьи и аннотация даются и на английском языке. После аннотации дать ключевые слова на русском и английском языках.
Электронная копия	Текстовый редактор Microsoft Word (версии 6.0, 7.0, 97). В имени файла указывается фамилия автора.
Параметры страницы	Формат А4. Поля: правое - 15 мм, левое - 25 мм, верхнее, нижнее - 20 мм.
Форматирование основного текста	С нумерацией страниц. Абзацный отступ - 5 мм. Интервал - полуторный.
Гарнитура шрифта	Times New Roman. Обычный размер кегля - 14 пт. Список литературы и аннотация - 12 пт.
Объем статьи (ориентировочно)	Кратких сообщений – до 3 с., статей на соискание ученой степени кандидата наук – 7-12 с., на соискание ученой степени доктора наук – 8-16 с.
Сведения об авторах	Указываются фамилия, имя, отчество (полностью), ученая степень, звание, должность и место работы, адрес с почтовым индексом, телефоны/факсы, e-mail (на русском и английском языках)

- Список литературы – все работы необходимо пронумеровать, в тексте ссылки на литературу оформлять в квадратных скобках.

- Материалы, не соответствующие предъявленным требованиям, к рассмотрению не принимаются.

- Решение о публикации статьи принимается редакцией «Вестника БГУ». Корректурa авторам не высылается, присланные материалы не возвращаются.

- Статьи принимаются в течение учебного года.

- Допустима публикация статей на английском языке, сведения об авторах, название и аннотацию которых необходимо перевести на русский язык.

- Формат журнала 60x84 1/8.

Стоимость обработки 1 с. (формата А4) для преподавателей БГУ составляет 200 р., для остальных – 400 р. Для аспирантов – бесплатно.

Адрес: 670000, г. Улан-Удэ, ул. Смолина, 24 а, Издательство БГУ.

Тел. (301-2)-21-95-57. Факс (301-2)-21-05-88

Оплата производится при получении счета от бухгалтерии БГУ.