

Карельский научный центр  
Российской академии наук  
Институт прикладных  
математических исследований

Karelian Research Center of  
the Russian Academy of Sciences  
Institute of Applied  
Mathematical Research



ТРУДЫ ИНСТИТУТА ПРИКЛАДНЫХ  
МАТЕМАТИЧЕСКИХ ИССЛЕДОВАНИЙ

**Методы математического моделирования  
и информационные технологии**

Выпуск 7

Петрозаводск  
2006

PROCEEDINGS OF THE INSTITUTE OF  
APPLIED MATHEMATICAL RESEARCH

Methods of mathematical modeling  
and information technologies

Volume 7

Petrozavodsk  
2006

**Редакционная коллегия:**

В.В.Мазалов (главный редактор), д.ф.-м.н., проф.  
Ю.Л.Павлов, д.ф.-м.н., проф.  
Ю.В.Заика, д.ф.-м.н., проф.  
В.Т.Вдовицын, к.ф.-м.н., доцент  
В.А.Лебедев, к.э.н., ст.н.с.  
А.В.Соколов, к.ф.-м.н., доцент  
Т.П.Тихомирова (ответственный за выпуск), к.т.н.  
доцент

ISBN 5-9274-0257-7

© Карельский научный центр РАН, 2006  
© Институт прикладных математических  
исследований КарНЦ РАН, 2006

## **Содержание**

Предисловие. . . . .	9
----------------------	---

### **Раздел I. МЕТОДЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ**

Абакумов Ю. Г., Банин М. В. О множествах сходимости, гарантируемой теоремами тестового типа. . . . .	13
Иванов А. П., Остов Ю. Я. Двойственность и принцип расширения в задаче динамики полета. . . . .	26
Попов В. В. Моделирование процессов тепломассопереноса в гидридном аккумуляторе водорода. . . . .	35
Попов В. В., Семин П. А. Реализация и использование задач моделирования взаимодействия водорода с твердым телом в специальном программном комплексе. . . . .	47
Стafeев С. В. О параметрической идентифицируемости модели факторного анализа с зависимыми факторами и остатками. . . . .	63
Тарарин В. М., Чеплюкова И. А. О действии случайного $s$ -отображения конечной булевой решетки. . . . .	80
Фалько А. А. Игра наилучшего выбора с возможностью отказа от предложения и перераспределением вероятностей. . . . .	87
Хвостянская Е. В. О возникновении гигантского $\varphi$ -примитивного идеала случайного автоморфизма $\varphi$ конечной булевой решетки. . . . .	95
Чернов И. А. Сходимость сеточных аппроксимаций решения краевой задачи с нелинейным граничными условием и подвижной границей. . . . .	97

### **Раздел II. ИНФОРМАТИКА**

Бородина А. В., Морозов Е. В. Доверительное оценивание вероятности переполнения буфера на основе ускоренного регенеративного моделирования для системы $M/M/1$ . . . . .	125
Киселева Н. С. Об одной модели случайного графа "Интернет" типа. . . . .	136
Кукин В. Д. Композитный эволюционный алгоритм для потоковой задачи Штейнера. . . . .	143

<i>Лазутина А. А. Оптимальное управление тремя стеками в случае параллельного выполнения операций.....</i>	154
<i>Печников А. А. Задача рационального размещения ссылок в регламентируемой локализованной системе интернет -ресурсов.....</i>	176
<i>Родченкова Н. И. Разностная схема для моделирования ТДС-спектра дегидрирования металлов с учетом теплопоглощения и сжатия.....</i>	183
<i>Реттиева А.Н., Родионов А.В. Моделирование экономических отношений в лесном комплексе Республики Карелия.....</i>	199
<i>Щербак М. П. Применение параллельных вычислений для решения некоторых задач механики твердого деформируемого тела.....</i>	206
<i>Abstracts .....</i>	216

## Table of Contents

Preface .....	9
---------------	---

### Section I. METHODS OF MATHEMATICAL MODELLING

<i>Abakumov Y.G., Banin M.V.</i> On sets of convergences guaranteed by test type theorems. ....	13
<i>Ivanov A.P., Ostov Yu.Ya.</i> Duality and principle of expansion in the problem of dynamics of flight. ....	26
<i>Popov V.V.</i> Modelling of heat-mass transfer processes in hydride accumulator of hydrogen. ....	35
<i>Popov V.V., Semin P.A.</i> Implementation and use of modelling problems of hydrogen interaction with solids in special software. ....	47
<i>Stafeev S.V.</i> On the parameter identifiability of a factor analysis model with correlated residuals and correlated factors. ....	63
<i>Tararin V.M., Cheplyukova I.A.</i> On the action of a random s-mapping of a finite Boolean lattice. ....	80
<i>Falko A.A.</i> A best-choice game with the possibility of an applicant refusing an offer and with redistribution of probabilities. ....	87
<i>Khvorostyanskaya E.V.</i> On emergence of the giant $\varphi$ -primitive ideal of a random automorphism $\varphi$ of a finite Boolean lattice. ....	95
<i>Chernov I.A.</i> Convergence of the lattice approximations of the solution to the boundary-value problem with nonlinear boundary condition and moving bound. ....	97

### Section II. COMPUTER SCIENCE

<i>Borodina A.V., Morozov E.V.</i> Overflow probability estimation with accelerated regeneration simulation in the queue $M/M/1$ . ....	125
<i>Kiseleva N.S.</i> On an "Internet" type random graph model. ....	136
<i>Kukin V.D.</i> A composite evolutionary algorithm for solving the weighted Steiner tree problem. ....	143

<i>Lazutina A.A.</i> On the problem of optimal control of three parallel stacks in bounded space.....	154
<i>Pechnikov A.A.</i> The problem of rational allocation of links in a controlled and localized Internet-resource system.....	176
<i>Rodchenkova N.I.</i> A differences scheme for modelling the TDS-spectrum of dehydriding of metals with heat absorption and a size reduction effect.....	183
<i>Rettieva A.N., Rodionov A.V.</i> The model of dynamic game of forest resources management problem is constructed.....	199
<i>Shcherbak M.P.</i> Parallel computing in application to analysis of some problems in mechanics.....	206
Abstracts .....	216

## **Предисловие**

Настоящий сборник статей продолжает серию работ Института прикладных математических исследований Карельского научного центра РАН. Он издается в юбилейный для Карельского научного центра РАН год, КарНЦ исполняется 60 лет.

Раздел, касающийся методов математического моделирования, содержит ряд работ, представленных на Российско-Скандинавском симпозиуме "Теория вероятностей и прикладная вероятность", который состоялся в Петрозаводске 26-31 августа 2006 г., а также на Российско-Финской школе-семинаре "Динамические игры и многокритериальная оптимизация", которая состоялась в Петрозаводске 2-7 сентября 2006г. Проведение этих научных мероприятий было поддержано Российским фондом фундаментальных исследований (проекты 06-01-10090, 06.01.10123) и Президиумом РАН по программе "Поддержка молодых ученых".

Раздел, посвященный информационным технологиям, содержит ряд работ, которые докладывались на международной конференции "Развитие вычислительной техники в России и странах бывшего СССР: история и перспективы", которая состоялась в Петрозаводске 3-7 июля 2006 г.

Директор Института прикладных  
математических исследований,  
д.ф.-м.н., проф.

В.В.Мазалов



**Раздел I.** Методы математического  
моделирования

**Section I.** Methods of Mathematical  
Modelling



УДК 517.51

## О МНОЖЕСТВАХ СХОДИМОСТИ, ГАРАНТИРУЕМОЙ ТЕОРЕМАМИ ТЕСТОВОГО ТИПА

Ю. Г. АБАКУМОВ, В. Г. БАНИН

Рассмотрены множества в пространстве  $C[a, b]$ , на которых из сходимости последовательности линейных операторов некоторых классов на тестовом множестве следует сходимость этой последовательности и на всем множестве. Показано, что расширить эти множества с сохранением сходимости нельзя. Рассмотрена общая схема для последовательности линейных функционалов в банаховом пространстве.

В 1953 г. П.П. Коровкин [1] доказал теорему о сходимости последовательности линейных положительных операторов  $L_n : C[a, b] \rightarrow C[a, b]$  (операторы положительные, если из условия  $f(t) \geq 0$  следует  $L_n(f(t), x) \geq 0$  для любых  $x \in [a, b]$ ), согласно которой выполняется следующее утверждение:

$$(\|L_n(t^i, x) - x^i\| \rightarrow 0 \text{ для } i \in \{0, 1, 2\}) \Rightarrow (\|L_n(f(t), x) - f(x)\| \rightarrow 0 \text{ для любой } f(t) \in C[a, b])$$

Здесь  $\|\cdot\|$  - чебышевская норма, т.е.  $\|f(x)\| = \max_{x \in [a, b]} |f(x)|$

По терминологии, используемой в работе Л.Г. Лабскера [2], пространство  $C[a, b]$  является  $(S_0^0\{I\})$  - замыканием Коровкина для множества  $\{1, t, t^2\}$  (здесь  $S_0^0$  обозначает множество положительных операторов,  $I$  - тождественный оператор).

Мы будем использовать иную терминологию. В аналогичных ситуациях будем говорить, что теорема тестового типа гарантирует сходимость последовательностей операторов данного класса (положительных в теореме Коровкина) на элементах данного множества (пространства  $C[a, b]$  в случае выше названной теоремы).

Следует отметить, что на практике при получении аналогов теоремы Коровкина встречались ситуации, когда без дополнительного условия равномерной ограниченности по  $n$  норм  $\|L_n\|$  не удавалось по сходимости на тестовом множестве гарантировать сходимость на всем пространстве  $C[a, b]$  (см. [3], [4], [5]).

В заметке [5], в достаточно частной ситуации, было выяснено, что отсутствие гарантии сходимости для всех  $f(t) \in C[a, b]$  имеет принципиальный характер.

Вопрос об определении множества гарантируемой сходимости (или, по терминологии [2], замыкания Коровкина) можно поставить следующим образом (ограничиваясь пока пространством  $C[a, b]$ ). Пусть имеется некоторый класс операторов  $\sigma_0$  и множество  $F \subset C[a, b]$ . Определить множество  $\Phi \subset C[a, b]$ , для которого:

**(a):** из  $L_n \in \sigma_0$ ,  $\|L_n(\varphi(t), x) - \varphi(x)\| \rightarrow 0$  для всех  $\varphi(t) \in F$  следует

$$\|L_n(f(t), x) - f(x)\| \rightarrow 0 \text{ для всех } f(t) \in \Phi$$

**(b):** какова бы ни была  $\psi(t) \in C[a, b]$ ,  $\psi(t) \notin \Phi$ , существует последовательность  $\Lambda_n \in \sigma_0$  такая, что

$$\|\Lambda_n(\varphi(t), x) - \varphi(x)\| \rightarrow 0 \text{ для всех } \varphi(t) \in F, \quad \overline{\lim}_{n \rightarrow \infty} \|\Lambda_n(\psi(t), x) - \psi(x)\| > 0$$

Для установления того, что данное множество  $\Phi$  удовлетворяет условию (a), имеются хорошо опробованные методики. Наиболее популярная из них основана на схемах, являющихся аналогами схемы Коровкина. В [6], например, некоторые из таких схем рассмотрены в достаточно общем виде. С проверкой свойства (b) дело обстоит сложнее. Наша попытка поставить этот вопрос и предложить некоторые подходы к его решению (см. [7]), насколько нам известно, внимания не привлекла.

В предлагаемой статье мы изложим, в частности, результаты депонированной работы [7]. Основные идеи работы [7] сохранены, но в детали внесены значительные изменения. Рассмотрены примеры.

Прежде чем обращаться к конкретизациям, рассмотрим некоторую общую ситуацию.

Введем основные элементы нашего исследования.

Обозначим  $X$  — действительное банахово пространство,  $\mu$  — фиксированный линейный непрерывный функционал ( $\mu : X \rightarrow R$ ),  $Ker\mu$  — ядро  $\mu$ , то есть множество его нулей,  $K \subset Ker\mu$

— конус, то есть такое выпуклое замкнутое множество, которое вместе с любым своим элементом  $p$  содержит множество  $\bigcup_{\lambda \geq 0} \lambda p (\lambda \in R)$ . Обозначим далее,  $K^*$  — сопряженный к  $K$  конус, то есть множество линейных непрерывных функционалов такое, что  $(\eta \in K^*) \Leftrightarrow (\forall p \in K, \eta(p) \geq 0)$ . Считаем, что зафиксировано некоторое конечномерное пространство  $F \subset X$ , причем, множества  $(F \cap K) \setminus \{0\}$  (здесь  $0$  — нуль пространства  $X$ ) и  $F \setminus \text{Ker } \mu$  непусты.

Для каждого элемента  $\bar{p} \in K$  обозначим  $L(\bar{p}) \subset \text{Ker } \mu$  такое множество, что  $p \in L(\bar{p})$  тогда и только тогда, когда существует действительное число  $c > 0$  (зависящее от  $p$ ) такое, что

$$c\bar{p} - p \in K, -c\bar{p} - p \in -K. \quad (1)$$

Обозначим далее

$$L(F, K) = \text{lin}\left(\bigcup_{\bar{p} \in F \cap K} L(\bar{p})\right) \cup F,$$

где  $\text{lin } Z$  — это линейная оболочка множества  $Z$ .

**Теорема 1.** Пусть  $\mu_n \in K^*$  и  $\forall g \in F \mu_n(g) \rightarrow \mu(g)$ . Тогда для всех  $p \in L(K, F)$  выполняется  $\mu_n(p) \rightarrow \mu(p)$ .

ДОКАЗАТЕЛЬСТВО. Пусть  $p \in L(K, F)$ . Тогда найдутся такие элементы  $p_1, p_2$ , что  $p_1 \in F$ ,  $p_2 \in \text{Ker } \mu$ ,  $p = p_1 + p_2$ ,  $\bar{p} \in F \cap K$ ,  $p_2 \in L(\bar{p})$ .

Следовательно, должны иметь место включения, аналогичные (1)  $c\bar{p} - p_2 \in K$ ,  $-c\bar{p} - p_2 \in -K$ , из которых, учитывая, что  $\mu_n \in K^*$ , получим  $c\mu_n(\bar{p}) \geq \mu_n(p_2)$ ,  $-c\mu_n(\bar{p}) \leq \mu_n(p_2)$  или  $|\mu_n(p_2)| \leq c|\mu_n(\bar{p})| \rightarrow 0$ .

Так как  $\mu(p) = \mu(p_1) + \mu(p_2) = \mu(p_1)$ ,  $\mu_n(p_1) \rightarrow \mu(p_1)$ ,  $\mu_n(p_2) \rightarrow 0$ , получаем  $\mu_n(p) = \mu_n(p_1) + \mu_n(p_2) \rightarrow \mu(p_1) = \mu(p)$ .

Теорема доказана.

Далее мы предполагаем, что  $K$  не содержит прямую (то есть, если  $p \neq 0$  и  $p \in K$ , то  $-p \notin K$ ) и установим, что в этом случае условие  $p \in L(K, F)$  в теореме 1 неустранимо.

Предварительно рассмотрим более детально условия принадлежности элемента  $p \in \text{Ker } \mu$  множеству  $L(\bar{p})$  для некоторого данного  $\bar{p} \in K$  и элемента  $p \in X$  множеству  $L(K, F)$ .

**Предложение 1.** Если для  $p \in \text{Ker}\mu$  и  $\bar{p} \in K$  найдутся положительные числа  $c_1$  и  $c_2$ , такие, что  $c_1\bar{p} - p \in K$ ,  $-c_2\bar{p} - p \in -K$ , то  $p \in L(\bar{p})$ .

ДОКАЗАТЕЛЬСТВО. Очевидно, для любых  $g \in K$ ,  $\lambda > 0$   $g + \lambda\bar{p} \in K$ , так как сумма двух элементов, каждый из которых принадлежит конусу, также принадлежит конусу.

Точно также ( $g \in -K, \lambda > 0$ )  $\Rightarrow g - \lambda\bar{p} \in -K$ .

Обозначим  $c = \max(c_1, c_2)$ . Если  $c = c_1$ , то  $-c\bar{p} - p = -(c - c_2)\bar{p} - c_2\bar{p} - p \in -K$ , так как  $-(c - c_2)\bar{p} \in -K$ ,  $-c_2\bar{p} - p \in -K$ . Согласно (1), в этом случае  $p \in L(\bar{p})$ . Симметричный случай  $c = c_2$  рассматривается аналогично.

Предложение доказано.

Обозначим  $\Pi(p, \bar{p})$  прямую, проходящую через  $p$  и  $\bar{p}$  (предполагая, что  $p \neq \bar{p}$ ), то есть

$$\Pi(p, \bar{p}) = \bigcup_{\lambda \in (-\infty, \infty)} \lambda p + (1 - \lambda)\bar{p}.$$

**Предложение 2.** Для  $p \in \text{Ker}\mu$ ,  $\bar{p} \in K$ , ( $p \neq \bar{p}$ ) включение  $p \in L(\bar{p})$  имеет место тогда и только тогда, когда  $\bar{p}$  является внутренней точкой относительно множества  $K \cap \Pi(p, \bar{p})$ .

ДОКАЗАТЕЛЬСТВО. Если  $\bar{p}$  является внутренней точкой относительно  $K \cap \Pi(p, \bar{p})$ , то найдутся  $\lambda_1, \lambda_2$ , причем,  $\lambda_1 > 0$ ,  $\lambda_2 < 0$ , такие, что  $\lambda p + (1 - \lambda)\bar{p} \in K$  при  $\lambda \in [\lambda_2, \lambda_1]$ .

Возьмем  $\tilde{\lambda}$  так, чтобы  $0 < \tilde{\lambda} < \min(1, \lambda_1)$ . Имеем  $\tilde{\lambda}p + (1 - \tilde{\lambda})\bar{p} \in K$ . Отсюда

$$-\tilde{\lambda}^{-1}(\tilde{\lambda}p + (1 - \tilde{\lambda})\bar{p}) = -p - \tilde{\lambda}^{-1}(1 - \tilde{\lambda})\bar{p} \in -K$$

или, если обозначить  $c_2 = \tilde{\lambda}^{-1}(1 - \tilde{\lambda})$ , включение можно записать в виде  $-c_2\bar{p} - p \in -K$ . Возьмем теперь

$$\tilde{\lambda} \in (\lambda_2; 0).$$

Тогда  $\tilde{\lambda}p + (1 - \tilde{\lambda})\bar{p} \in K$ . Отсюда (так как  $\tilde{\lambda} < 0$ ),

$$-\tilde{\lambda}^{-1}(\tilde{\lambda}p + (1 - \tilde{\lambda})\bar{p}) = -p + \tilde{\lambda}^{-1}(\tilde{\lambda} - 1)\bar{p} \in K.$$

Обозначив  $c_1 = \tilde{\lambda}^{-1}(\tilde{\lambda} - 1)$ , запишем  $c_1\bar{p} - p \in K$ . Согласно предложению 1,  $p \in L(\bar{p})$ .

Итак, доказано, что если  $\bar{p}$  является внутренней точкой относительно  $K \cap \Pi(p, \bar{p})$ , то  $p \in L(\bar{p})$ .

Докажем обратное утверждение.

Пусть  $p \in L(\bar{p})$ . Тогда найдется  $c_0 > 0$  такое, что  $-p + c_0\bar{p} \in K$ ,  $-p - c_0\bar{p} \in -K$ . Для любого  $c > c_0$  имеем  $(c - c_0)\bar{p} \in K$ ,  $(c_0 - c)\bar{p} \in -K$ , следовательно,  $-p + c\bar{p} \in K$ ,  $-p - c\bar{p} \in -K$ . Отсюда следует, что в (1) можно положить  $c > 1$ .

Из  $-p + c\bar{p} \in K$  получим

$$(c - 1)^{-1}(-p + c\bar{p}) = -(c - 1)^{-1}p + (1 - (-c - 1)^{-1})\bar{p} \in K.$$

Из  $-p - c\bar{p} \in -K$  получим  $p + c\bar{p} \in K$ . Отсюда

$$(c + 1)^{-1}(p + c\bar{p}) = (c + 1)^{-1}p + (1 - (c + 1)^{-1})\bar{p} \in K.$$

В силу выпуклости  $K$   $\lambda p + (1 - \lambda)\bar{p} \in K$  при  $-(c - 1)^{-1} \leq \lambda \leq (c + 1)^{-1}$ . Это означает, что  $\bar{p}$  является внутренней точкой относительно  $K \cap \Pi(p, \bar{p})$ .

Предложение 2 доказано.

В дальнейшем обозначаем  $H = F \cap \text{Ker}\mu$ ,  $\bar{K} = K \cap H$ . Для  $p \in \text{Ker}\mu$  обозначим  $H(p) = \text{lin}(H \cup \{p\})$ ,  $K(p) = K \cap H(p)$ .

Для того, чтобы  $p \in L(K, F)$ , необходимо и достаточно, чтобы в  $H(p)$  нашелся элемент  $\tilde{p} \notin H$ , а в  $\bar{K}$  элемент  $\bar{p}$  такие, что  $\tilde{p} \in L(\bar{p})$ . Согласно предложению 2, это равносильно тому, что  $\bar{p}$  является внутренней точкой множества  $K(p) \cap \Pi(\tilde{p}, \bar{p})$ . Элемента  $\tilde{p}$  с указанными выше свойствами не найдется в том и только в том случае, когда  $\bar{K} = K(p)$  или  $\bar{K}$  является гранью  $K(p)$  (говорят, что конус  $K_1$  является гранью конуса  $K_2$ , если  $K_1$  правильная часть  $K_2$  и  $K_1$  есть пересечение  $K_2$  с какой-либо опорной к  $K_2$  гиперплоскостью).

Сказанное выше можно сформулировать в виде следующего утверждения.

**Предложение 3.** Элемент  $p \in \text{Ker}\mu$  не принадлежит  $L(K, F)$  тогда и только тогда, когда  $\bar{K} = K(p)$  или  $\bar{K}$  является гранью  $K(p)$ .

Основной результат данной статьи представляет следующая теорема.

**Теорема 2.** Пусть  $p \notin L(K, F)$ . Тогда существует последовательность  $\{\mu_n\}$ ,  $\mu_n \in K^*$  такая, что

$$\overline{\lim_{n \rightarrow \infty}} |\mu_n(p) - \mu(p)| \neq 0, \text{ но } \forall z \in F \mu_n(z) \rightarrow \mu(z)$$

ДОКАЗАТЕЛЬСТВО. Сначала докажем для случая  $p \in Ker\mu$ .

Согласно предложению 3, из того, что  $p \notin L(K, F)$  следует, что  $\bar{K} = K(p)$  либо  $\bar{K}$  является гранью  $K(p)$ .

Мы будем строить функционалы  $\mu_n$ , о которых говорится в условии теоремы 2. Сначала определим их в пространстве  $H$ , потом доопределим на пространство  $H(p)$ , затем на  $lin(F \cup \{p\})$  и, наконец, докажем возможность продолжения  $\mu_n$  на все пространство  $X$ .

Выберем какую-либо последовательность действительных положительных чисел  $\varepsilon_n \downarrow 0$  и элемент  $g$ , являющийся внутренним относительно  $\bar{K}$ .

Пусть  $G$  — какая-либо гиперплоскость в пространстве  $H$ , такая, что  $G \cap K = \{0\}$  (где 0 — нуль пространства  $X$ ). Функционал  $\mu_n$  определяется на  $H$  условиями:  $\mu_n(g) = \varepsilon_n$ ,  $\forall r \in b \mu_n(r) = 0$ . Очевидно, что для любого  $y \in H$   $\mu_n(y) \rightarrow 0$ . Действительно, можно представить  $y = r + cg$ , где  $r \in G$ ,  $c$  — действительное число, не зависящее от  $n$ . Тогда  $\mu_n(y) = \mu_n(r) + c\mu_n(g) = c\varepsilon_n \rightarrow 0$ .

Распространим  $\mu_n$  на  $H(p)$ . Выберем какую-либо последовательность действительных положительных чисел  $A_n$  с тем условием, чтобы  $\lim_{n \rightarrow \infty} A_n > 0$  (не исключено, что  $A_n \rightarrow \infty$ ). Так как  $\bar{K}$  является гранью  $K(p)$ , выполняется по крайней мере одно из двух условий: (i)  $[p, g] \cap K(p) = \{g\}$ , (ii)  $[2g-p, g] \cap K(p) = \{g\}$ . (Здесь, как обычно,  $[p, g]$  обозначает замкнутый отрезок, т.е. множество элементов вида  $\lambda p + (1-\lambda)g$  при  $\lambda \in [0, 1]$ ).

Если выполнено (i), то положим  $p' = 2g-p$ , в противном случае положим  $p' = p$ . Доопределим  $\mu_n$  на  $H(p)$ , приняв  $\mu_n(p') = A_n$ .

Заметим, что аннулятор функционала  $\mu_n$  в пространстве  $H(p)$  проходит через точку  $\lambda_n p' + (1-\lambda_n)g$  при  $\lambda_n = -\frac{\varepsilon_n}{A_n - \varepsilon_n}$ , и, следовательно, находится на некотором расстоянии от  $g \in \bar{K}$ .

Далее,  $\mu_n(p) = A_n$  при условии (ii) или (если мы положили  $p' = 2g-p$ )  $\mu_n(p) = 2\varepsilon_n - A_n$ . И в том, и в другом случае  $\lim_{n \rightarrow \infty} |\mu_n(p)| > 0$ .

Продолжим теперь функционалы  $\mu_n$  на пространство  $lin(F \cup \{p\})$ . Для этого выберем в  $lin F$  элементы  $z_i$ ,  $i = 1, \dots, s$  так, чтобы

они были линейно независимы, чтобы  $\text{lin}(\{z_i\} \cup H) = F$  и чтобы при любом  $i \in \{1, \dots, s\}$   $z_i \notin H$ . Положим  $\mu_n(z_i) = \mu(z_i)$ . Так как  $K(p) \subset H(p)$ , а все  $z_i \notin H(p)$ , положительность  $\mu_n$  на  $K(p)$  сохраняется.

Теперь докажем возможность продолжения  $\mu_n$  на  $X$  с сохранением положительности на  $K$ .

Для любого положительного, достаточно малого числа  $\delta$  обозначим  $K(g, \delta)$  коническую оболочку объединения  $K$  и замкнутого шара с центром  $g$  радиуса  $\delta$ . В силу того, что аннулятор  $\mu_n$  в  $H(p)$  (очевидно, и в  $\text{lin}(F \cup \{p\})$ ) находится на положительном расстоянии от  $p \in K$ , можно подобрать  $\delta_n$  так, что  $\mu_n$  останутся положительными на  $K(g, \delta_n) \cap \text{lin}(F \cup \{p\})$ . Так как конус  $K(g, \delta_n)$  телесный, то  $\mu_n$  можно продолжить на  $X$  с сохранением положительности на  $K(g, \delta_n)$ , а, следовательно, и на  $K$  (см. [8], стр. 394).

В силу построения, для любого  $y \in F$   $\mu_n(y) \rightarrow \mu(y)$ , однако,  $\mu_n(p)$  не стремится к  $\mu(p) = 0$ .

В случае  $p \notin L(K, F)$ ,  $\mu(p) = 0$  теорема доказана.

Если же  $p \in L(K, F)$ ,  $\mu(p) \neq 0$ , то можно представить  $p = p_1 + p_2$ , где  $p_1 \in F$ ,  $\mu(p_2) = 0$ . Тогда  $p_2 \notin L(K, F)$ , иначе бы  $p$  принадлежал  $L(K, F)$ . Как показано выше, существует последовательность  $\mu_n \in K^*$  такая, что  $\mu_n(y) \rightarrow \mu(y)$  для  $y \in F$ ,  $\lim_{n \rightarrow \infty} |\mu_n(p_2)| > 0$ . Но тогда

$$\lim_{n \rightarrow \infty} \mu_n(p) = \lim_{n \rightarrow \infty} \mu_n(p_1) + \lim_{n \rightarrow \infty} \mu_n(p_2) \neq \mu(p_1) = \mu(p).$$

Теорема 2 доказана.

Рассмотрим некоторые конкретизации.

**Пример 1.** Некоторые операторы класса  $S_m$ .

Заметим, что для понимания изложенного ниже общего определения операторов класса  $S_m$ , данного, например, в [2], не требуется.

Пусть каждому  $x \in [a, b]$  поставлено в соответствие множество  $T(x) = \{x_i\}_{i=1}^{m(x)}$ , при этом  $x \in T(x)$ ,  $m(x) > 1$  при любом  $x$  и существует конечный максимум  $m = \max_{x \in [a, b]} m(x)$ . Кроме того, определена функция  $e(x)$ ,  $e : [a, b] \rightarrow \{-1; 1\}$ . Обозначим  $w_x(t) = e(x) \prod_{x_i \in T(x)} (t - x_i)$ .

Для каждого  $x \in [a, b]$  определим конус  $K_x$  следующим образом:  $K_x$  — минимальный по включению конус, такой, что  $(\text{sign } f(t) = w_x(t)) \Rightarrow f(t) \in K_x$ .

Полагаем далее, что существует  $\Delta > 0$  такое, что  $\forall x \in [a, b]$   $\Delta(x) \geq \Delta$ , где  $\Delta(x) = \min_{t, \tau \in T(x), t \neq \tau} |t - \tau|$ .

Будем рассматривать последовательности линейных операторов, удовлетворяющих условию:  $\forall x \in [a, b]$  из  $f(t) \in K_x \Rightarrow L_n(f(t), x) \geq 0$ .

**Предложение 4.** Пусть  $L_n$  — линейные операторы, удовлетворяющие приведенным выше условиям и, кроме того,  $\sup_x |L_n(t^i, x) - x_i| \rightarrow 0$  при  $i = 0, 1, \dots, m$ . Тогда

$$\forall f(t) \in Lip1 : \lim_{n \rightarrow \infty} \sup_x |L_n(f(t), x) - f(x)| = 0.$$

**Доказательство.** Зафиксируем произвольно  $x \in [a, b]$  и рассмотрим последовательность функционалов  $L_n(f(t), x)$ , которые, в силу условия, положительны на конусе  $K_x$ .

Пусть, далее, зафиксирована функция  $f(t) \in LipM1$ .

Для того, чтобы применить теорему 1, обозначим соответствие элементов рассматриваемой конкретной ситуации элементам общей схемы.

Так, конусу  $K$  соответствует конус  $K_x$ , пространству  $F$  — множество  $A_m$  — алгебраических полиномов степени  $m$ , элементу  $p$  — функция  $f(t)$ . В качестве функционала  $\mu$  выступает  $\delta$  — функция точки  $x$ .

Покажем, что (в принятых ранее обозначениях)  $f(t) \in L(K_x, A_m)$ , при этом элементу  $p_1$  соответствует  $g(t)$  — интерполяционный полином Лагранжа функции  $f(t)$  по узлам  $x_i \in T(x)$ , элементу  $p_2$  — разность  $f(t) - g(t)$ , а в качестве  $\bar{p}$  можно взять  $w_x(t)$ .

Действительно,  $g(t) \in A_m$ , а так как  $x \in T(x)$ , то  $f(x) - g(x) = 0$ , то есть  $f(t) - g(t) \in Ker\mu$ . Далее,  $w_x(t) \in K_x$  и, как известно,  $|f(t) - g(t)| = |f(t, x_1, \dots, x_m)| \cdot |w_x(t)|$ , где  $f(t, x_1, \dots, x_m)$  — разделенная разность. Таким образом, если существует конечная верхняя грань  $\sup_{t, x \in [a, b]} |f(t, x_1, \dots, x_m)|$ , то ее или любое число большее нее можно взять в качестве множителя  $c$  в формуле (1), причем, для любого  $x \in [a, b]$ .

Можно убедиться, что все величины вида  $f(t, \tau_1, \dots, \tau_k)$  по модулю равномерно ограничены сверху (для фиксированной  $f(t)$ ). Здесь  $k \leq m(x)$ ,  $\{\tau_i\}_{i=1}^k \subset T(x)$ , все элементы множества  $\{t, \tau_1, \dots, \tau_k\}$  различны.

Действительно, при  $k = 1$  получаем  $|f(t, \tau_1)| = \frac{|f(\tau_1) - f(t)|}{|\tau_1 - t|} \leq M$ .  
При  $k = 2$

$$|f(t, \tau_1, \tau_2)| = |f(\tau_1, t, \tau_2)| = \frac{1}{|\tau_2 - \tau_1|} |f(t, \tau_2) - f(t, \tau_1)| \leq 2M\Delta^{-1}.$$

Далее, если неравенство

$$|f(t, \tau_1, \dots, \tau_{k-1})| \leq 2^{k-2}M\Delta^{-k+2}$$

выполнено для любого набора  $\{t, \tau_1, \dots, \tau_{k-1}\}$ , удовлетворяющего приведенным выше ограничениям, то для  $\{t, \tau_1, \dots, \tau_k\}$  получаем

$$\begin{aligned} & |f(t, \tau_1, \tau_2, \dots, \tau_k)| = |f(\tau_1, t, \tau_2, \dots, \tau_k)| = \\ & = \frac{1}{|\tau_k - \tau_1|} |f(t, \tau_2, \dots, \tau_k) - f(\tau_1, t, \tau_2, \dots, \tau_{k-1})|. \end{aligned}$$

Отсюда

$$\begin{aligned} & |f(t, \tau_1, \dots, \tau_k)| \leq \frac{1}{|\tau_k - \tau_1|} (|f(t, \tau_2, \dots, \tau_k)| + |f(t, \tau_1, \dots, \tau_{k-1})|) \leq \\ & \leq \Delta^{-1} \cdot 2 \cdot 2^{k-2}M\Delta^{-k+2} = 2^{k-1}M\Delta^{-k+1}. \end{aligned}$$

Итак, неравенство

$$|f(t, \tau_1, \dots, \tau_k)| 2^{k-1}M\Delta^{-k+1}$$

выполняется для всех  $k = 1, \dots, m(x)$ .

Таким образом, если положить  $c = \max_x 2^{m(x)-1}M\Delta^{-m(x)+1}$ , то включения  $c w_x(t) - (f(t) - g(t)) \in K_x$ ,  $-c w_x(t) - (f(t) - g(t)) \in -K_x$  имеют место при любом  $x \in [a, b]$ .

Из очевидного равенства  $L_n(f(t), x) = L_n(g(t), x) + L_n(f(t) - g(t), x)$  и полученного в ходе доказательства теоремы 1 неравенства  $|\mu_n(p_2)| \leq c|\mu_n(\bar{p})|$ , получим

$$|L_n(f(t), x) - f(x)| \leq |L_n(g(t), x) - g(x)| - cL_n(w_x(t), x). \quad (2)$$

(Модуль в последнем слагаемом правой части неравенства (2) не поставлен, так как  $w_x(t) \in K_x$ ). Здесь было использовано равенство  $f(x) = g(x)$ . Для константы  $c$  в неравенстве (2) можно взять отмеченное выше значение  $c = \max_x 2^{m(x)-1}M\Delta^{-m(x)+1}$ .

Так как  $g(t)$  и  $w_x(t)$  — полиномы степени  $m$  и  $w_x(x) = 0$ , то из неравенства (2) и условий предложения 4 следует поточечная сходимость  $L_n(f(t), x)$  к  $f(x)$ . Для доказательства равномерной сходимости следует установить равномерное стремление к нулю

$L_n(w_x(t), x)$  и  $|L_n(g(t), x) - g(x)|$ , что не очевидно, так как  $g(t)$  и  $w_x(t)$  зависят от  $x$ .

Функции  $g(t)$  и  $w_x(t)$  являются алгебраическими полиномами степени не выше  $m$ , то есть можно представить

$$w_x(t) = \sum_{i=0}^{m(x)} a_i t^i, \quad g(t) = \sum_{i=0}^{m(x)-1} b_i t^i.$$

Так как  $|w_x(t)| = |\prod_{i=1}^{m(x)} (t - x_i)|$ , то очевидно, что  $|a_i|$  равномерно по  $x$  ограничены. Положим  $A = \sup_{x,i} |a_i|$ . Далее, так как

$g(t) = f(x_1) + f(x_1, x_2)(t - x_1) + \dots + f(x_1, \dots, x_{m(x)})(t - x_1) \cdots (t - x_{m(x)-1})$  и, как мы установили выше,  $|f(x_1, \dots, x_k)|$  равномерно ограничены, то равномерно по  $x$  ограничены и  $|b_i|$ . Обозначим  $B = \sup_{x,i} |b_i|$ . Если обозначить  $\gamma_n = \max_{0 \leq i \leq m} \sup_x |L_n(t^i, x) - x^i|$ , то из (2) получим оценку

$$|L_n(f(t), x) - f(x)| \leq ((m+1)cA + mB)\gamma_n. \quad (3)$$

Так как  $\gamma_n$  и все константы в (3) не зависят от  $x$ , получим

$$\lim_{n \rightarrow \infty} \sup_x |L_n(f(t), x) - f(x)| = 0.$$

Предложение 4 доказано.

**Предложение 5.** Пусть  $\psi(t) \in C[a, b]$  такова, что существует точка  $\xi \in [a, b]$  такая, что выполняется хотя бы одно из равенств  $|\psi'_+(\xi)| = \infty$  или  $|\psi'_-(\xi)| = \infty$ . Тогда существует последовательность линейных операторов  $L_n(f(t), x)$ , удовлетворяющих условиям предложения 4, для которых

$$\lim_{n \rightarrow \infty} \sup_x |L_n(t^i, x) - x^i| = 0$$

при  $i = 0, 1, \dots, m$ , но

$$\lim_{n \rightarrow \infty} \sup_x |L_n(\psi(t), x) - \psi(x)| > 0.$$

**ДОКАЗАТЕЛЬСТВО.** Рассмотрим функцию  $\Theta(t) = \psi(t) - g_\psi(t)$ , где  $g_\psi(t)$  – интерполяционный многочлен Лагранжа функции  $\psi(t)$  по узлам, принадлежащим  $T(\xi)$ .

Если для какой-либо последовательности линейных операторов выполняются условия:

$$L_n(t^i, \xi) \rightarrow \xi^i, \quad i = 0, 1, \dots, m, \quad \lim_{n \rightarrow \infty} |L_n(\Theta(t), \xi)| > 0$$

то, очевидно,

$$\lim_{n \rightarrow \infty} |L_n(\psi(t), \xi) - \psi(\xi)| > 0.$$

Таким образом, для доказательства предложения 5 (в качестве первого шага) мы установим, что существует последовательность линейных функционалов  $\mu_n \in K_\xi^*$  такая, что

$$\mu_n(t^i) \rightarrow \xi^i, i = 0, 1, \dots, m, \quad \lim_{n \rightarrow \infty} |\mu_n(\Theta(t))| > 0. \quad (4)$$

Для того, чтобы утверждать существование последовательности  $\mu_n$  с указанными выше свойствами, нам будет достаточно сослаться на теорему 2, если мы докажем, что  $\Theta(t) \notin L(K_\xi, A_m)$ . Для доказательства этого применим предложение 3. Действительно, в данном случае  $\bar{K}_\xi$  состоит из полиномов вида  $w_\xi(t) \cdot s'(t)$ , где  $s'(t)$  – неотрицательный на  $[a, b]$  алгебраический полином такой, что степень  $w_\xi(t) \cdot s'(t)$  не выше  $m$ . Если  $\Theta_0(t) \in \text{lin}(\bar{K}_\xi \cup \{\Theta(t)\})$ ,  $\Theta(t) \notin \bar{K}_\xi$ ,  $\eta(t) \in \bar{K}_\xi$ , то  $\eta(t)$  не является внутренней точкой (относительно прямой  $\Pi(\Theta_0(t), \eta(t))$ ) множества  $K_\xi \cap \Pi(\Theta_0(t), \eta(t))$ . В самом деле, одна из односторонних производных функции  $\Theta_0(t)$  в точке  $\xi$  не ограничена, следовательно, эта же односторонняя производная не ограничена и у любой из функций вида  $\lambda\Theta_0(t) + (1 - \lambda)\eta(t)$ , если  $\lambda \neq 0$ . Кроме того, при изменении знака  $\lambda$  меняет знак и упоминаемая односторонняя производная (с  $+\infty$  на  $-\infty$  или наоборот). Очевидно, что конусы  $K_\xi$  не могут принадлежать одновременно две функции, имеющие в точке  $\xi$  односторонние производные разных знаков. Итак,  $\eta(t)$  не является внутренней относительно прямой  $\Pi(\Theta_0(t), \eta(t))$  точкой множества  $K_\xi \cap \Pi(\Theta_0(t), \eta(t))$ . Следовательно,  $\bar{K}_\xi$  является гранью  $K_\xi(\Theta_0(t))$ . Применяя предложение 3 и теорему 2, заключаем, что существует последовательность функционалов  $\mu_n \in K_\xi^*$ , удовлетворяющая условиям (4).

Если мы теперь определим последовательность операторов  $L_n$  с помощью равенств:  $L_n(f(t), \xi) = \mu_n(f(t))$ ,  $L_n(f(t), x) = f(x)$  при  $x \in [a, b] \setminus \{\xi\}$ , то, очевидно, последовательность удовлетворяет условиям:

$$\lim_{n \rightarrow \infty} \sup_x |L_n(t^i, x) - x^i| = 0$$

при

$$i = 0, \dots, m, \quad \lim_{n \rightarrow \infty} \sup_x |L_n(\psi(t), x) - \psi(x)| > 0.$$

Предложение 5 доказано.

**Пример 2.** Операторы классов  $\beta_m(\varepsilon)$  и  $\beta_m^-(\varepsilon)$ .

Определение операторов этих классов дано в [8]. Оператор  $L : C[a, b] \rightarrow C[a, b]$  принадлежит классу  $\beta_m(\varepsilon)$  (классу  $\beta_m^-(\varepsilon)$ ), если  $\forall x \in [a, b]$  из того, что

$$f(t) \in C^m[a, b], f^{(i)}(x) = 0$$

при

$$i = 0, 1, \dots, m-1, f^{(m)}(t) > 0 (f^{(m)}(t) < 0)$$

и

$$\frac{P_{\max}}{P_{\min}} \leqslant \frac{1 + \varepsilon}{1 - \varepsilon}, \text{ где } P_{\max} = \max_{t \in [a, b]} |f^{(m)}(t)|, P_{\min} = \min_{t \in [a, b]} |f^{(m)}(t)|$$

следует  $L(f(t), x) \geqslant 0$ .

В [8] доказана теорема об условиях сходимости последовательностей операторов классов  $\beta_m(\varepsilon)$  и  $\beta_m^-(\varepsilon)$ , а также установлено, что операторы  $\tilde{M}_n(f(t), x)$ , являющиеся алгебраическими аналогами операторов, построенных В.А. Баскаковым [9], обладают следующим свойством: для любого  $\varepsilon \in (0; 1)$  можно найти  $n_0$  такое, что при  $n > n_0$   $\tilde{M}_n \in \beta_4^-(\varepsilon)$ .

Приведём формулировку Теоремы ([8], стр. 75): Пусть  $f(t) \in C^m[a, b]$  и, при этом,  $|f^{(m)}(t)| \leqslant \mu > 0$ ,  $L_n \in \beta_m(\varepsilon)$  (или  $L_n \in \beta_m^-(\varepsilon)$ ), тогда

$$|L_n(f(t), x) - \sum_{k=0}^{m-1} \frac{f^{(k)}(x)}{k!} L_n((t-x)^k, x)| \leqslant \frac{\mu \varepsilon^{-1}}{m!} |L_n((t-x)^m, x)|.$$

Для каждого  $x$   $K_x$  является замыканием множества функций вида  $\lambda \varphi(t), \lambda \geqslant 0, \varphi(t) \in C^m[a, b], 1 - \varepsilon \leqslant \varphi^{(m)}(t) \leqslant 1 + \varepsilon$  (в случае класса  $\beta_m^-(\varepsilon)$  должно выполняться  $-1 - \varepsilon \leqslant \varphi^{(m)}(t) \leqslant -1 + \varepsilon$ ).

По той же схеме, что и в примере 1, можно установить, что условия приведенной выше теоремы из [8] не гарантируют сходимость  $L_n(\psi(t), x)$  к  $\psi(x)$  при условии

$$\lim_{n \rightarrow \infty} \sup_x |L_n(t^i, x) - x^i| = 0 \text{ для } i = 0, 1, \dots, m$$

если хотя бы для одной точки  $\xi \in (a, b)$  выполняется  $|\psi^{(m)}(\xi)| = \infty$ .

В самом деле, так как для любого элемента  $\varphi(t) \in K_\xi$  найдется  $c > 0$  такое, что  $|\varphi(t)| \leqslant c|t - \xi|^m$ , то, очевидно, если  $|\psi^{(m)}(\xi)| = \infty$ , то  $\psi(t) + P_m(t) \notin K_\xi$ , где  $P_m(t)$  — полином степени не выше

*m.* Тогда  $\tilde{K}_\xi = K_\xi(\psi)$ . Достаточно сослаться на предложение 3 и теорему 2.

### Литература

1. Коровкин П.П. О сходимости линейных положительных операторов в пространстве непрерывных функций // Докл. АН СССР. 1953, т. 90, № 6, с. 961-964.
2. Лабскер Л.Г. Тестовые множества для приближения операторами и функционалами класса  $S_m$ . М.: МИПКХП, 1990.
3. Баскаков В.А. Обобщение некоторых теорем П.П. Коровкина о положительных операторах // Мат.заметки. 1973, № 13, с. 785-794.
4. Абакумов Ю.Г., Лампе Е.И. Об одной теореме типа теоремы П.П. Коровкина // Применение функ. анализа в теории прибл. Калинин: КГУ, 1974, вып. 2, с. 3-7.
5. Абакумов Ю.Г. Аналог теоремы П.П. Коровкина для одного класса линейных операторов // Применение функ. анализа в теории прибл. Калинин: КГУ, 1975, вып. 6, с. 3-4.
6. Абакумов Ю.Г., Забелина Н.А., Шестакова О.Н. О последовательностях линейных функционалов и некоторых операторах класса  $S_{2m}$  // Сиб. матем. журнал. 2000, т. 41, № 2, с. 247-252.
7. Абакумов Ю.Г., Банин В.Г. Модификации теорем Климова, Красносельского и Лифшица и исследование аппроксимативных свойств линейных операторов. Ч.1 // Читинский гос. пед. ин-т. Чита, 1987. Деп. в ВИНТИ 17.11.87, № 8094.
8. Абакумов Ю.Г., Банин В.Г. Об операторах класса  $\beta_m(\varepsilon)$  и об алгебраических аналогах тригонометрических операторов Баскакова // Диф. уравнения и аналитическая теория. Чита: ЧитГТУ, 1999, с. 75-80.
9. Баскаков В.А. Об одном методе построения операторов класса  $S_{2m}$  // Теория функций и приближений. Интерполирование по Лагранжу. Саратов, 1984, с.19-25.

УДК 517.9

## ДВОЙСТВЕННОСТЬ И ПРИНЦИП РАСШИРЕНИЯ В ЗАДАЧЕ ДИНАМИКИ ПОЛЕТА

А. П. Иванов, Ю. Я. Остов

Рассматривается задача максимизации кинетической энергии летательного аппарата в конечной точке траектории. Предложен метод построения субоптимального управления.

Задача оптимального управления движением летательного аппарата (ЛА) в атмосфере является одной из практически важных задач внешней баллистики и теории управления. Сложность проблемы обусловлена трудоемкостью решения, получаемого численными методами на основе принципа максимума Л. С. Понtryгина [1,2]. Помимо этого, экстремальное (оптимальное) управление, найденное как результат решения краевой задачи на основе принципа максимума в его классической формулировке, является программным управлением, и при наличии всякого рода возмущений оказывается неэффективным, т. е. не обеспечивает оптимум заданного критерия качества.

Поэтому, если говорить о прикладном аспекте решения задачи, то нет необходимости в «точном» решении задачи оптимизации: достаточно построить управление, при котором значение оптимизируемого функционала отличается от его оптимального значения не более, чем на заданную величину  $\epsilon$ . Такое управление будем называть субоптимальным (оптимальным с точностью до  $\epsilon$  в смысле значения оптимизируемого функционала). Практически это достигается за счет упрощения модели, описывающей реальный управляемый процесс и заданной обычно в виде

системы дифференциальных уравнений. Упрощение этой системы уравнений возможно за счет рационального выбора фазового пространства, независимой переменной интегрирования, замены исходных связей и функционала эквивалентными.

Однако в большинстве случаев, даже с учетом специфики прикладной задачи, такое упрощение модели не дает желаемого результата в виде синтезированного управления. Поэтому дальнейшее упрощение модели (второй этап) осуществляется на основе принципа расширения, приводящего к вырожденной задаче оптимального управления. И, наконец, третий этап решения задачи (при восстановленной связи, исключенной на втором этапе) приводит к желаемому результату в виде синтеза (управления с обратной связью) [3,5].

Предлагаемую методику можно проиллюстрировать на примере: оптимизировать траекторию продольного движения центра масс (ЦМ) ЛА, совершающего полет из начальной точки атмосферного пространства в заданную конечную точку на поверхности Земли. Критерием оптимальности траектории (управления) является максимум кинетической энергии ЛА в конечной точке траектории.

Движение ЦМ ЛА описывается следующими уравнениями [4]:

$$\begin{aligned} \frac{dV}{dt} &= -\frac{X}{m} - g \sin \theta, & V(0) &= V_0, \\ \frac{d\theta}{dt} &= \frac{Y}{mV} - \frac{g \cos \theta}{V}, & \theta(0) &= \theta_0, \\ \frac{dH}{dt} &= V \sin \theta, & H(0) &= H_0, \\ \frac{d\xi}{dt} &= \beta \bar{\rho} V, & \xi(0) &= \xi_0, \end{aligned} \quad (1)$$

где  $V(t)$  — модуль скорости,  $\theta(t)$  — угол наклона траектории к горизонту,  $H(t)$  — высота над поверхностью Земли,  $\xi(t)$  — безразмерная «взвешенная» длина траектории,  $t \in [0, T]$ ,  $m$  — масса ЛА,  $g$  — ускорение свободного падения,  $X = (C_{x0} + C_{xi}\alpha^2)\rho V^2 S/2$  — лобовое сопротивление ЛА,  $Y = C_y^\alpha \alpha \rho V^2 S/2$  — подъемная сила ЛА,  $\rho = \rho_0^* \exp(-\beta H)$  — плотность атмосферы на высоте  $H$ ,  $\bar{\rho} \equiv \rho/\rho_0^*$ ,  $\alpha$  — угол атаки (управление).

В модели  $C_{x0}$ ,  $C_{xi}$ ,  $C_y^\alpha$ ,  $S$ ,  $m$ ,  $\rho_0^*$ ,  $\beta$ ,  $g$  — заданные константы.

Ставится задача: максимизировать кинетическую энергию

$$J(\alpha) = \varphi(V(T)) = mV^2(T)/2 \quad (2)$$

при заданных значениях высоты и «взвешенной» длины траектории полета в конечный момент  $T$ , т. е. при заданных  $H_T, \xi_T$

$$H(T) - H_T = 0, \quad (3)$$

$$\xi(T) - \xi_T = 0. \quad (4)$$

Угол  $\theta(T)$  и конечный момент времени  $T$  не фиксированы, что не влияет на общность подхода к решению задачи.

На первом этапе исходная система (1) преобразуется к виду:

$$\begin{aligned} \frac{dV_l}{d\xi} &= -\left(\bar{c}_{x0} + \frac{\bar{c}_y^2}{2d}\right)V_l - \bar{c}_y V_h, & V_l(\xi_0) &= V_{l0}, \\ \frac{dV_h}{d\xi} &= \bar{c}_y V_l - \left(\bar{c}_{x0} + \frac{\bar{c}_y^2}{2d}\right)V_h - \frac{\bar{g}}{\bar{\rho}\sqrt{V_l^2 + V_h^2}}, & V_h(\xi_0) &= V_{h0}, \\ \frac{d\bar{\rho}}{d\xi} &= -\frac{V_h}{\sqrt{V_l^2 + V_h^2}}, & \bar{\rho}(\xi_0) &= \bar{\rho}_0, \end{aligned} \quad (5)$$

где

$$\bar{\rho} = \frac{\rho}{\rho_0^*} = \exp(-\bar{h}), \quad \bar{h} = \beta H, \quad \bar{g} = \frac{g}{\beta},$$

$$\bar{c}_{x0} = \frac{C_{x0}\rho_0^*S}{2m\beta}, \quad d = \frac{(C_y^\alpha)^2 \rho_0^* S}{4m\beta C_{xi}}, \quad \bar{c}_y = \frac{(C_y^\alpha)\rho_0^* S \alpha}{2m\beta}, \quad (6)$$

$V_l, V_h$  — проекции скорости  $\bar{V}$  ЦМ ЛА соответственно на горизонтальную и вертикальную оси инерциальной системы.

Функционал (2) и ограничения (3), (4) перепишутся в виде:

$$\tilde{J}(u) = \tilde{\varphi}((V_l^2 + V_h^2)(\xi_T)) = V^2(\xi_T), \quad (7)$$

$$\bar{\rho}(\xi_T) - \bar{\rho}_T = 0, \quad (8)$$

$$\xi \in [\xi_0, \xi_T], \quad \xi_0 \stackrel{\Delta}{=} 0. \quad (9)$$

Как следует из приведенных соотношений, ЛА обладает параболической полярой, т. е. коэффициент лобового сопротивления связан с коэффициентом аэродинамической подъемной силы соотношением  $\bar{c}_x = c_{x0} + \bar{c}_y^2/2d$ , где  $c_{x0}$  и  $d$  — положительные константы. Из последнего соотношения следует  $y = 2\sqrt{x-a}/b$ , где приняты обозначения  $y = \bar{c}_y$ ,  $x = \bar{c}_x$ ,  $a = \bar{c}_{x0}$ ,  $b = \sqrt{2/d}$ .

Обозначим  $G(x) = 2\sqrt{x-a}/b$ . Пусть  $x_k$  — коэффициент лобового сопротивления. Касательная к кривой  $G(x)$  в точке  $x = x_k$  представляется в виде:

$$y = \frac{\partial G}{\partial x} \Big|_{x=x_k} (x - x_k) + G(x_k)$$

или

$$y = \frac{(x - x_k)}{b\sqrt{x_k - a}} + 2\frac{\sqrt{x_k - a}}{b}.$$

Обозначим точку пересечения прямой с осью  $Ox$  через  $z$ , т. е.

$$0 = \frac{(z - x_k)}{b\sqrt{x_k - a}} + 2\frac{\sqrt{x_k - a}}{b}.$$

Откуда следует  $x_k + z = 2a$ ,  $a = \text{const}$ . С учетом последних двух соотношений уравнение касательной принимает вид

$$y = \frac{(x - z)}{b\sqrt{a - z}}.$$

Точке параболы ( $\bar{c}_x = x$ ,  $\bar{c}_y = y$ ) сопоставлена прямая (в пространственном случае — гиперплоскость, которая соответствует преобразованию Лежандра применительно к параболоиду).

Теперь система уравнений (5) перепишется в виде

$$\begin{aligned} \frac{dV_l}{d\xi} &= -\bar{c}_x V_l - \frac{(\bar{c}_x - z)}{b\sqrt{a - z}} V_h, & V_l(\xi_0) &= V_{l0}, \\ \frac{dV_h}{d\xi} &= \frac{(\bar{c}_x - z)}{b\sqrt{a - z}} V_l - \bar{c}_x V_h - \frac{\bar{g}}{\bar{\rho}\sqrt{V_l^2 + V_h^2}}, & V_h(\xi_0) &= V_{h0}, \\ \frac{d\bar{\rho}}{d\xi} &= -\frac{V_h}{\sqrt{V_l^2 + V_h^2}}, & \bar{\rho}(\xi_0) &= \bar{\rho}_0. \end{aligned} \quad (10)$$

При переходе к новой независимой переменной  $\bar{\rho}$ :

$$\begin{aligned}\frac{dV_l}{d\bar{\rho}} &= \frac{\bar{c}_x V_l \sqrt{V_l^2 + V_h^2}}{V_h} + \frac{(\bar{c}_x - z) \sqrt{V_l^2 + V_h^2}}{b\sqrt{a-z}}, \\ \frac{dV_h}{d\bar{\rho}} &= -\frac{(\bar{c}_x - z)V_l \sqrt{V_l^2 + V_h^2}}{bV_h \sqrt{a-z}} + \bar{c}_x \sqrt{V_l^2 + V_h^2} + \frac{\bar{g}}{\bar{\rho}V_h}, \\ \frac{d\xi}{d\bar{\rho}} &= -\frac{\sqrt{V_l^2 + V_h^2}}{V_h}.\end{aligned}\quad (11)$$

Гамильтониан системы (11) и сопряженная система уравнений:

$$H = \Psi_0 \frac{dV_l}{d\bar{\rho}} + \Psi_1 \frac{dV_h}{d\bar{\rho}} + \Psi_2 \frac{d\xi}{d\bar{\rho}}, \quad (12)$$

$$\begin{aligned}\frac{d\Psi_0}{d\bar{\rho}} &= -\frac{\partial H}{\partial V_l} = -\Psi_0 \Phi_0 - \Psi_1 \Phi_1 - \Psi_2 \Phi_2, \\ \frac{d\Psi_1}{d\rho} &= -\frac{\partial H}{\partial V_h} = -\Psi_0 \Phi_3 - \Psi_1 \Phi_4 - \Psi_2 \Phi_5, \\ \frac{d\Psi_2}{d\rho} &= -\frac{\partial H}{\partial \xi} = 0,\end{aligned}\quad (13)$$

где

$$\begin{aligned}\Phi_0 &= \frac{\bar{c}_x V}{V_h} + \frac{\bar{c}_x V_l^2}{V_h V} + \frac{(\bar{c}_x - z)V_l}{bV\sqrt{a-z}}, \quad \Phi_2 = -\frac{V_l}{VV_h}, \\ \Phi_1 &= -\frac{(\bar{c}_x - z)V}{bV_h\sqrt{a-z}} - \frac{(\bar{c}_x - z)V_l^2}{bVV_h\sqrt{a-z}} + \frac{\bar{c}_x V_l}{V}, \\ \Phi_3 &= \frac{-\bar{c}_x V_l^3}{VV_h^2} + \frac{(\bar{c}_x - z)V_h}{bV\sqrt{a-z}}, \\ \Phi_4 &= \frac{(\bar{c}_x - z)V_l^3}{bVV_h^2\sqrt{a-z}} + \frac{\bar{c}_x V_h}{V} - \frac{\bar{g}}{\bar{\rho}V_h^2}, \quad \Phi_5 = \frac{V_l^2}{VV_h^2}.\end{aligned}$$

Так как  $V = \sqrt{V_l^2 + V_h^2} > 0$ , условие стационарности гамильтониана относительно управления  $\bar{c}_x$  можно представить в виде:

$$F_1 = \frac{1}{V} \frac{\partial H}{\partial \bar{c}_x} = \Psi_0 \left( \frac{V_l}{V_h} + \frac{1}{b\sqrt{a-z}} \right) + \Psi_1 \left( -\frac{V_l}{bV_h\sqrt{a-z}} + 1 \right) = 0. \quad (14)$$

Условие стационарности гамильтониана относительно функции  $z(\bar{\rho})$  при ограничении  $(a - z) > 0$ :

$$-\frac{1}{V} \frac{\partial H}{\partial z} 2b(a - z)^{3/2} = \left( \Psi_0 - \Psi_1 \frac{V_l}{V_h} \right) (2a - \bar{c}_x - z) = 0.$$

Первый сомножитель в правой части этого выражения отличен от нуля (иначе  $\Psi_0 = \Psi_1 = 0$ ), второй сомножитель равен нулю.

Гамильтониан на экстремали является функцией  $\bar{\rho}$ :  $H = H(\bar{\rho})$ ,  $H(\bar{\rho}) = -C \sqrt{V_l^2 + V_h^2}/V_h$ ,  $C = \text{const} > 0$ . Поэтому

$$F_2 = \left( \frac{\Psi_0 V_l}{V_h} + \Psi_1 \right) z + \frac{C - \Psi_2}{V_h} + \frac{\Psi_1 \bar{g}}{\sqrt{V_l^2 + V_h^2} \bar{\rho} V_h} = 0. \quad (15)$$

Продифференцировав интегралы

$$F_1(X(\bar{\rho}), z(\bar{\rho})) = 0 \quad \text{и} \quad F_2(X(\bar{\rho}), z(\bar{\rho}), \bar{\rho}) = 0$$

по  $\bar{\rho}$ , где  $X = (V_l, V_h, \Psi_0, \Psi_1)$  в силу системы (11), (13), и исключив переменную  $dz/d\bar{\rho}$ , получим интеграл  $F_3 = 0$  в виде:

$$F_3 = a_{11} a_{22} - a_{12} a_{21} = 0,$$

где

$$a_{11} = \frac{\partial F_1}{\partial X} \frac{dX}{d\bar{\rho}}, \quad a_{22} = \frac{\partial F_2}{\partial z}, \quad a_{12} = \frac{\partial F_1}{\partial z}, \quad a_{21} = \frac{\partial F_2}{\partial X} \frac{dX}{d\bar{\rho}} + \frac{\partial F_2}{\partial \bar{\rho}}.$$

В развернутом виде интеграл  $F_3 = 0$  можно представить так:

$$F_3 = \frac{\bar{g} \Psi_1}{\bar{\rho}^2 V} + \left( \frac{-2 \sqrt{a - z} V_l V}{b V_h^2} + \frac{\bar{g} V_l^2}{\bar{\rho} V^2 V_h^2} \right) C = 0. \quad (16)$$

Получаем следующие выражения:

$$\Psi_0 = \frac{V (V_l - b Ir V_h) \bar{\rho} (\Psi_2 - C)}{z V^3 \bar{\rho} + \bar{g} b Ir V_l + V_h \bar{g}}, \quad \Psi_1 = \frac{(b Ir V_l + V_h) (\Psi_2 - C) \bar{\rho} V}{z V^3 \bar{\rho} + \bar{g} b Ir V_l + V_h \bar{g}},$$

где  $Ir = \sqrt{a - z}$ . Подставив далее  $\Psi_1$  в интеграл (16), получим уравнение 3-ей степени относительно переменной  $\bar{c}_x$ :

$$P_1^2 (\bar{c}_x - a) - P_0^2 = 0. \quad (17)$$

Без ограничения общности можно принять  $\Psi_2 = C_1 + C$ , а  $C = 1$ , тогда  $P_0$  и  $P_1$  примут следующий вид:

$$P_0 = -3 \frac{\cos^2 \theta \bar{c}_x}{\sin^2 \theta} + \frac{\bar{g} \cos^2 \theta}{\sin \theta V^2 \bar{\rho}} + \frac{\sin \theta C_1}{\bar{\rho}} + 4 \frac{\cos^2 \theta a}{\sin^2 \theta},$$

$$P_1 = 2 \frac{\bar{\rho} V^2 \cos \theta \bar{c}_x}{\sin^2 \theta b \bar{g}} + \frac{\bar{g} b \cos^3 \theta}{\sin^2 \theta V^2 \bar{\rho}} - 2 \frac{\cos \theta}{\sin \theta b} - 4 \frac{\bar{\rho} V^2 \cos \theta a}{\sin^2 \theta \bar{g} b} + \frac{b \cos \theta C_1}{\bar{\rho}},$$

где  $\theta = \text{Arctan}(V_h/V_l)$ ,  $V = \sqrt{V_l^2 + V_h^2}$ ,  $a = \bar{c}_{x0}$ . Результаты численного моделирования, приведенные ниже, подтверждают эффективность рассмотренной методики. Синтезированное по упрощенной модели управление  $\bar{c}_x$  находилось как корень уравнения (17) при текущих значениях фазовых переменных  $\theta$ ,  $\bar{\rho}$ ,  $V$  и пересчитывалось в угол атаки  $\alpha$ , который в исходной системе (1) является управлением. При этом константы  $b$  и  $C_1$ , от которых зависят коэффициенты полинома (17), должны быть подобраны так, чтобы были выполнены заданные граничные условия. Система уравнений (1) интегрировалась методом Рунге-Кutta (4-го порядка точности) с шагом  $h_u = 0.02$  при следующих значениях параметров ЛА и атмосферы:  $C_{x0} = 0.1931$ ,  $C_{xi} = 5.88$ ,  $C_y^\alpha = 0.8548046$ ,  $m = 422$ ;  $S = 0.159$ ,  $g = 9.81$ .

Вариант 1:  $\rho_0^* = 2.047$ ,  $\beta = 1.5682E - 04$ .

Вариант 2:  $\rho_0^* = 1.225$ ,  $\beta = 1.4280E - 04$ .

Начальная точка траектории в обоих вариантах одна и та же:  $\xi(0) = 0$ ,  $H(0) = 24054.5$ ,  $V(0) = 1088.31$ ,  $\theta(0) = -0.54113$ . Терминальные ограничения (3), (4) в обоих вариантах следующие:  $H(T) - 0 = 0$ ,  $\xi(T) - 1.42500 = 0$ . Результаты счета этих вариантов:

Параметр	Упрощенная модель	Полная модель
$b$	13.367120	---
$C_1$	0.422322464	---
$T$	37.88	37.80
$\xi(T)$	1.42500	1.42500
$H(T)$	-0.00001	0.00000
$V(T)$	700.84701	700.8499
$\theta(T)$	-0.844160	-0.844283

Параметр	Упрощенная модель	Полная модель
$b$	6.633430	---
$C_1$	0.459130	---
$T$	36.18	36.162
$\xi(T)$	1.42500	1.42500
$H(T)$	-0.00001	0.0000
$V(T)$	858.9485	858.96097
$\theta(T)$	-0.82950	-0.826503

Третий вариант отличается от первого начальной точкой траектории:  $\xi(0) = 0.0$ ,  $H(0) = 30000.0$ ,  $V(0) = 1200.0$ ,  $\theta(0) = -0.4$ , а также ограничением на конечную точку траектории:  $\xi(T) - 1.6500 = 0$ . Остальные параметры соответствуют варианту 1.

В четвертом варианте параметры атмосферы  $\rho_0^*$  и  $\beta$  как во втором варианте, остальные параметры соответствуют варианту 3. Результаты счета вариантов 3 и 4 приведены соответственно в таблицах:

Параметр	Упрощенная модель	Полная модель
$b$	4.87190	---
$C_1$	0.7479072	--
$T$	50.93	51.01
$\xi(T)$	1.6500	1.6500
$H(T)$	-0.00001	0.00000
$V(T)$	695.8655	695.93738
$\theta(T)$	-0.72382	-0.72557847

Параметр	Упрощенная модель	Полная модель
$b$	6.242810	---
$C_1$	0.8558571	--
$T$	48.45	48.50
$\xi(T)$	1.6500	1.6500
$H(T)$	0.00002	0.00000
$V(T)$	887.99959	888.0763
$\theta(T)$	-0.71095	-0.7129783

Примечание: под полной моделью следует понимать традиционную схему решения задачи (1)–(4) на основе принципа максимума Л.С.Понtryгина в его классической формулировке (с интегрированием сопряженной системы).

При управлении, синтезированном по упрощенной модели, система интегрируется примерно в 1,8 раза быстрее, чем при использовании традиционной схемы решения задачи (1)–(4), т. к. отпадает необходимость интегрировать сопряженную систему (13).

Следует заметить, что представление (16) интеграла  $F_3 = 0$  является не единственным, т. е. можно получить лучшее приближение оптимального значения функционала (2), а последнее уравнение системы (1) можно заменить другим соотношением, например, уравнением нормированной горизонтальной дальности  $dl/dt = \beta V \cos \theta$  или уравнением «взвешенной» горизонтальной дальности  $d\sigma/dt = \beta \bar{\rho} V \cos \theta$ .

## Литература

1. Понtryгин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М.: Наука, 1961.
2. Кирин Н.Е. Методы последовательных оценок в задачах оптимизации управляемых систем. Л.: АГУ, 1975.
3. Кротов В.Ф., Букреев В.З., Гурман В.И. Новые методы вариационного исчисления в динамике полета. М.: Наука, 1969.
4. Лебедев А.А., Чернобровкин Л.С. Динамика полета. М.: Наука, 1973.
5. Иванов А.П., Остов Ю.Я. Аналитический метод решения задачи субоптимального управления полетом // Методы математического моделирования и информационные технологии. Труды ИПМИ КарНЦ РАН. Вып. 4. Петрозаводск, 2003, с. 83–91.

УДК 519.6, 539.2

**МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ТЕПЛОМАССОПЕРЕНОСА  
В ГИДРИДНОМ АККУМУЛЯТОРЕ ВОДОРОДА**

**В. В. ПОПОВ**

Рассматривается одномерная модель процессов тепло-  
массопереноса в гидридном аккумуляторе водорода. Учи-  
тывается зависимость коэффициентов модели от темпе-  
ратуры и концентрации, а также протекание химиче-  
ских реакций образования и разложения гидрида.

Работа выполнена при поддержке "Фонда содействия  
отечественной науке".

**1. ВВЕДЕНИЕ**

В настоящее время в связи с перспективами водородной энергетики актуальной является проблема хранения и транспортировки водорода. В качестве перспективного способа решения этой проблемы рассматривается использование гидридных накопителей (аккумуляторов) водорода. Ввиду большого разнообразия возможных конструкций, в статье не рассматриваются модели конкретных аккумуляторов, а рассмотрена лишь одномерная математическая модель, описывающая основные физико-химические процессы. Модель может быть полезна для построения трехмерных моделей конкретных конструкций аккумуляторов, что в свою очередь позволит проводить расчет характеристик, выбирать оптимальные параметры, ставить и решать задачи оптимального управления их работой.

## 2. ВЫВОД УРАВНЕНИЙ ТЕПЛОМАССОПЕРЕНОСА

Рассмотрим работу гидридного аккумулятора водорода. В простейшем случае аккумулятор представляет собой некоторую емкость с трубками для подвода и отвода водорода, в которой находится порошок гидридообразующего металла. Внутрь емкости помещены элементы для охлаждения и нагрева порошка, поскольку реакции образования и разложения гидридов являются экзо(эндо)термическими. В начальный момент времени в емкость подается водород под давлением. При определенной концентрации водорода и температуре (достаточно низкой) начинается реакция гидридообразования с выделением тепла. После реакции водород оказывается химически связанным в гидриде металла. В таком виде его безопасно можно хранить и транспортировать. При необходимости получить водород, гидрид в аккумуляторе нагревают до температуры разложения. Реакция разложения гидрида является эндотермической, поэтому для поддержания реакции гидрид следует подогревать.

Для построения модели идеализируем конструкцию аккумулятора следующим образом. Рассмотрим полый цилиндр достаточно большого внутреннего диаметра  $d$ , имеющий дно только с одной стороны (рис. 1). Будем считать, что цилиндр заполнен порошком гидридообразующего металла, а нагревающие и охлаждающие элементы конструкции скрыты в стенках. Водород подается или отбирается с открытой стороны цилиндра. В силу такой идеализации можно рассматривать одномерную модель процессов, пренебрегая влиянием стенок цилиндра (при достаточно большом диаметре). Т.е. далее будем рассматривать некоторое твердое тело толщиной  $\ell$ , состоящее из порошинок гидридообразующего металла и имеющее пористость  $\eta(t, x)$ . При этом с одной стороны подается газообразный водород, с другой стороны тело соприкасается с некоторой стенкой. Основные физико-химические процессы, требующие рассмотрения в модели: диффузия (дисперсия) водорода в порах материала; теплоперенос в газе (водороде) и в материале (порошке); теплообмен соприкосновением и излучением между материалом и газом; образование гидрида с учетом поглощения водорода и выделения энергии или распад гидрида с выделением водорода и поглощением энергии. Диффузия водорода в материале не рассматривается ввиду

незначительности по сравнению с диффузией в порах. Все основные коэффициенты модели зависят от температуры, коэффициент диффузии зависит еще и от концентрации.



Рис. 1. Схема накопителя водорода

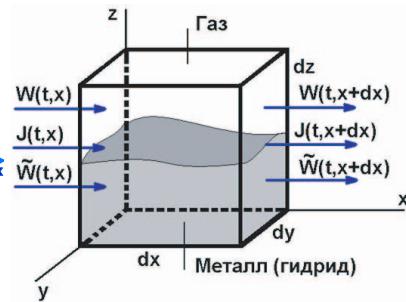


Рис. 2. Процессы в элементарном объеме

Для построения модели рассмотрим процессы, происходящие в элементарном объеме вещества (кубе) со сторонами  $dx, dy, dz$  (рис. 2) за время  $dt$ , и используем законы сохранения [1–3]. В этом объеме будем рассматривать диффузионный поток вещества  $J(t, x)$  и тепловые потоки в газе  $W(t, x)$  и в материале  $\tilde{W}(t, x)$  только вдоль оси  $x$ . Для определенности будем рассматривать процесс гидрирования, в случае дегидрирования модель выводится аналогичным образом. В дальнейшем будем использовать следующие обозначения (величины, относящиеся к порошку, имеют в обозначении надстрочный знак  $\sim$ ):  $c(t, x)$  — концентрация молекулярного водорода в порах ( $1/cm^3$ );  $T(t, x), \tilde{T}(t, x)$  — температура водорода и порошка ( $K$ );  $p(t, x)$  — давление водорода в порах ( $bar$ );  $q(t, x), \tilde{q}(t, x)$  — теплоемкости газа и порошка ( $J/gr\cdot K$ );  $\lambda(t, x), \tilde{\lambda}(t, x)$  — теплопроводности газа и порошка ( $J/s\cdot K\cdot cm$ );  $\rho(t, x), \tilde{\rho}(t, x)$  — плотности газа и порошка ( $gr/cm^3$ );  $D(t, x)$  — коэффициент диффузии (дисперсии) газа в материале ( $cm^2/s$ );  $\eta(t, x)$  — пористость порошка (отношение объема газа ко всему объему);  $a(t, x)$  — количество потребленного (выделенного) водорода при гидрировании (дегидрировании) ( $1/cm^3\cdot s$ );  $\mu$  — молярная масса водорода;  $N_A$  — число Авогадро.

**2.1. Вывод уравнения диффузии (дисперсии).** Количество молекул водорода принесенное диффузионным потоком и оставшееся в кубе это разница между диффузионными потоками через поверхности куба:

$$Q_1 = [J(t, x) - J(t, x + dx)] dy dz dt.$$

Внутри куба часть водорода поглощается в результате гидридообразования:

$$Q_2 = a(t, x) dx dy dz dt.$$

Изменение количества молекул водорода в кубе можно выразить с учетом пористости порошка иначе:

$$Q_3 = [c(t + dt, x)\eta(t + dt, x) - c(t, x)\eta(t, x)] dx dy dz.$$

Из закона сохранения вещества имеем  $Q_3 = Q_1 - Q_2$ . Подставляя выражение для диффузионного потока  $J(t, x) = -D(t, x)\partial c(t, x)/\partial x$ , получим:

$$\frac{\partial}{\partial t}(\eta(t, x)c(t, x)) = \frac{\partial}{\partial x}\left(D(t, x)\frac{\partial c(t, x)}{\partial x}\right) - a(t, x).$$

Преобразуем полученное уравнение к более привычному виду:

$$\frac{\partial c(t, x)}{\partial t} = \frac{1}{\eta(t, x)}\frac{\partial}{\partial x}\left(D(t, x)\frac{\partial c(t, x)}{\partial x}\right) - \frac{c(t, x)}{\eta(t, x)}\frac{\partial \eta(t, x)}{\partial t} - \frac{a(t, x)}{\eta(t, x)}. \quad (1)$$

**2.2. Вывод уравнения теплопроводности в газе.** Поток тепла изменяет внутреннюю энергию объема на величину:

$$E_1 = \left[W(t, x)\eta(t, x) - W(t, x + dx)\eta(t, x + dx)\right] dy dz dt.$$

Диффузионный поток приводит к изменению внутренней энергии объема на величину:

$$E_2 = \mu \left[J(t, x)q(t, x)T(t, x) - J(t, x + dx)q(t, x + dx)T(t, x + dx)\right] dy dz dt / N_A.$$

Энергия, полученная вследствие теплообмена с порошком:

$$E_3 = \sigma \left(\tilde{T}^4(t, x) - T^4(t, x)\right) + \alpha \left(\tilde{T}(t, x) - T(t, x)\right).$$

Тепловая энергия поглощенного порошком водорода:

$$E_4 = a(t, x)\mu q(t, x)T(t, x) / N_A.$$

Разогрев вещества в объеме выражается через температуру:

$$E_5 = \left[ \eta(t+dt, x)\rho(t+dt, x)q(t+dt, x)T(t+dt, x) - \right. \\ \left. - \eta(t, x)\rho(t, x)q(t, x)T(t, x) \right] dx dy dz.$$

Вследствии закона сохранения энергии имеем:  $E_5 = E_1 + E_2 + E_3 - E_4$ . Подставляя сюда  $\mu/N_A = \rho/c$  (получается из  $cV\mu/N_A = m$ , где  $V$  – объем;  $m$  – масса вещества),  $J(t, x) = -D(t, x)\partial c(t, x)/\partial x$ ,  $W(t, x) = -\lambda(t, x)\partial T(t, x)/\partial x$ , деля на  $dx dy dz dt$  и устремляя к нулю  $dx dt$ , получим (для краткости аргументы не указываются):

$$\eta_t \rho q T + \eta \rho_t q T + \eta \rho q_t T + \eta \rho q T_t = (\eta \lambda T_x)_x + \sigma(\tilde{T}^4 - T^4) + \alpha(\tilde{T} - T) + \\ + (\rho/c)[(Dc_x)_x q T + Dc_x q_x T + Dc_x q T_x - aq T].$$

Упростим полученное уравнение. Для этого перенесем первые три слагаемые из левой части в правую, используем уравнение диффузии (1) и учтем, что  $p_t = c_t \mu/N_A$ :

$$\eta \rho q T_t = (\eta \lambda T_x)_x + \sigma(\tilde{T}^4 - T^4) + \alpha(\tilde{T} - T) + (\rho/c)[Dc_x q_x T + Dc_x q T_x] - \eta \rho q_t T.$$

Теперь разделим на  $\eta \rho q$  и запишем в более привычном виде:

$$\frac{\partial T}{\partial t} = \frac{1}{\eta \rho q} \frac{\partial}{\partial x} \left( \eta \lambda \frac{\partial T}{\partial x} \right) + \frac{\sigma(\tilde{T}^4 - T^4) + \alpha(\tilde{T} - T)}{\eta \rho q} + \\ + \frac{DT}{c \eta q} \frac{\partial c}{\partial x} \frac{\partial q}{\partial x} + \frac{D}{c \eta} \frac{\partial c}{\partial x} \frac{\partial T}{\partial x} - \frac{T}{q} \frac{\partial q}{\partial t}. \quad (2)$$

**2.3. Вывод уравнения теплопроводности в порошке.** Вывод уравнения производится аналогично. Поток тепла изменяет внутреннюю энергию объема на величину:

$$\tilde{E}_1 = [\tilde{W}(t, x)(1 - \eta(t, x)) - \tilde{W}(t, x + dx)(1 - \eta(t, x + dx))] dy dz dt.$$

Энергия, отданная вследствие теплообмена с порошком:

$$\tilde{E}_2 = \sigma(\tilde{T}^4(t, x) - T^4(t, x)) + \alpha(\tilde{T}(t, x) - T(t, x)).$$

Тепловая энергия поглощенного водорода:

$$\tilde{E}_3 = a(t, x)\mu q(t, x)T(t, x)/N_A.$$

Энергия, выделившаяся в результате реакции гидрирования:

$$\tilde{E}_4 = a(t, x)H_E/N_A.$$

Разогрев вещества в объеме выражается через температуру:

$$\begin{aligned}\tilde{E}_5 = & \left[ (1 - \eta(t+dt, x))\tilde{\rho}(t+dt, x)\tilde{q}(t+dt, x)\tilde{T}(t+dt, x) - \right. \\ & \left. - (1 - \eta(t, x))\tilde{\rho}(t, x)\tilde{q}(t, x)\tilde{T}(t, x) \right] dx dy dz.\end{aligned}$$

Вследствии закона сохранения энергии имеем:  $\tilde{E}_5 = \tilde{E}_1 - \tilde{E}_2 + \tilde{E}_3 + \tilde{E}_4$ . Подставляя сюда  $\tilde{W}(t, x) = -\tilde{\lambda}(t, x)\partial\tilde{T}(t, x)/\partial x$ , деля на  $dx dy dz dt$  и устремляя к нулю  $dx dt$ , получим:

$$\begin{aligned}-\eta_t\tilde{\rho}\tilde{q}\tilde{T} + (1-\eta)\tilde{\rho}_t\tilde{q}\tilde{T} + (1-\eta)\tilde{\rho}\tilde{q}_t\tilde{T} + (1-\eta)\tilde{\rho}\tilde{q}\tilde{T}_t = \\ = ((1-\eta)\tilde{\lambda}\tilde{T}_x)_x - \sigma(\tilde{T}^4 - T^4) - \alpha(\tilde{T} - T) + a(\mu qT + H_E)/N_A.\end{aligned}$$

Перенесем первые три слагаемые из левой части в правую, разделим выражение на  $(1-\eta)\tilde{\rho}\tilde{q}$  и запишем в более привычном виде:

$$\begin{aligned}\frac{\partial\tilde{T}}{\partial t} = & \frac{1}{(1-\eta)\tilde{\rho}\tilde{q}} \frac{\partial}{\partial x} \left( (1-\eta)\tilde{\lambda} \frac{\partial\tilde{T}}{\partial x} \right) - \frac{\sigma(\tilde{T}^4 - T^4) + \alpha(\tilde{T} - T)}{(1-\eta)\tilde{\rho}\tilde{q}} + \\ & + \tilde{T} \left[ \frac{1}{(1-\eta)} \frac{\partial\eta}{\partial t} - \frac{1}{\tilde{q}} \frac{\partial\tilde{q}}{\partial t} - \frac{1}{\tilde{\rho}} \frac{\partial\tilde{\rho}}{\partial t} \right]. \quad (3)\end{aligned}$$

В уравнениях (2), (3) слагаемое  $\alpha(\tilde{T} - T)$  выражает теплообмен соприкосновением, а слагаемое  $\sigma(\tilde{T}^4 - T^4)$  выражает теплообмен излучением. Величины  $\alpha, \sigma$ , могут быть конкретизированы:  $\alpha = \alpha_0 S$ ,  $\sigma = \sigma_0 \beta S$  [4, 5]. Здесь  $\sigma_0 = 5.67 \cdot 10^{-8} J/s \cdot m^2 \cdot K^4$  – постоянная Стефана-Больцмана;  $S$  – площадь поверхности порошинок в элементарном объеме;  $\beta$  – коэффициент, характеризующий степень черноты поверхности;  $\alpha_0$  – коэффициент, характеризующий теплообмен соприкосновением.

### 3. ГРАНИЧНЫЕ И НАЧАЛЬНЫЕ УСЛОВИЯ

Будем полагать, что при гидрировании водород подается из некоторого фиксированного объема без поддержания заданного давления (концентрации), а при дегидрировании собирается в некотором объеме без отбора газа. В этом случае граничные условия для уравнений тепломассопереноса (1)–(3) для стороны, обращенной к газу, могут быть получены в интегральной и дифференциальной форме с использованием законов сохранения следующим образом (концентрация и температура в объеме  $V_0$  предполагается равной  $c(t, 0)$ ,  $T(t, 0)$ ).

Количество молекул водорода в объеме  $V_0$  в начальный момент времени равно  $N_A V_0 / V_{\mu,0}$  ( $V_{\mu,0}$  — мольный объем водорода при заданном начальном давлении  $p_0$ ), в произвольный момент времени  $V_0 c(t, 0)$ , разница этих количеств есть диффузионный поток через всю поверхность за прошедшее время  $S \int_0^t [D(\tau, 0) \partial c(\tau, 0) / \partial x] d\tau$ . Поэтому имеем:

$$V_0 c(t, 0) = \frac{N_A V_0}{V_{\mu,0}} + \int_0^t D(\tau, 0) \frac{\partial c(\tau, 0)}{\partial x} d\tau. \quad (4)$$

Тепловая энергия водорода в объеме  $V_0$  в начальный момент времени равна  $T(0, 0)q(0, 0)c(0, 0)V_0\mu/N_A$ , в произвольный момент времени  $T(t, 0)q(t, 0)c(t, 0)V_0\mu/N_A$ . Разность этих энергий это тепловые потоки через поверхность  $S \int_0^t [\eta(\tau, 0)\lambda(\tau, 0)\partial T(\tau, 0)/\partial x] d\tau$ ,  $S \int_0^t [(1 - \eta(\tau, 0))\tilde{\lambda}(\tau, 0)\partial \tilde{T}(\tau, 0)/\partial x] d\tau$  и энергия, унесенная вместе с диффузионным потоком:  $\mu S \int_0^t [q(\tau, 0)T(\tau, 0)D(\tau, 0)\partial c(\tau, 0)/\partial x] d\tau/N_A$ . Используя эти выражения, можно записать:

$$\begin{aligned} T(t, 0) \frac{q(t, 0)c(t, 0)V_0\mu}{N_A} &= T(0, 0) \frac{q(0, 0)c(0, 0)V_0\mu}{N_A} + \\ &+ S \int_0^t \eta(\tau, 0)\lambda(\tau, 0) \frac{\partial T(\tau, 0)}{\partial x} d\tau + S \int_0^t (1 - \eta(\tau, 0))\tilde{\lambda}(\tau, 0) \frac{\partial \tilde{T}(\tau, 0)}{\partial x} d\tau + \\ &+ \frac{\mu S}{N_A} \int_0^t q(\tau, 0)T(\tau, 0)D(\tau, 0) \frac{\partial c(\tau, 0)}{\partial x} d\tau. \end{aligned} \quad (5)$$

На границе порошка и водорода происходит теплообмен со-прикосновением и излучением, поэтому граничное условие будет определять тепловой поток:

$$(1 - \eta(t, 0))\tilde{\lambda}(t, 0) \frac{\partial \tilde{T}(t, 0)}{\partial x} = \sigma_1(\tilde{T}^4(t, 0) - T^4(t, 0)) + \alpha_1(\tilde{T}(t, 0) - T(t, 0)). \quad (6)$$

Здесь величины  $\sigma_1$ ,  $\alpha_1$  определяются аналогично  $\sigma$ ,  $\alpha$ . В случае подачи (откачки) водорода граничные условия модифицируются. Граничные условия (4), (5), записанные в интегральной форме, можно переписать в более привычной дифференциальной форме, т.е. в виде динамических граничных условий (дифференциальных уравнений для поверхностных величин).

Для стороны, граничащей со стенкой, граничные условия определяют потоки тепла и газа:

$$(1-\eta(t, \ell))\tilde{\lambda}(t, \ell) \frac{\partial \tilde{T}(t, \ell)}{\partial x} = F_1(t), \quad \eta(t, \ell)\lambda(t, \ell) \frac{\partial T(t, \ell)}{\partial x} = F_2(t), \quad (7)$$

$$D(t, \ell) \frac{\partial c(t, \ell)}{\partial x} = F_3(t). \quad (8)$$

В частности, тепловые потоки можно считать константами, а поток газа нулевым (нет диффузии водорода через стенки).

Начальными условиями для уравнений (1)–(3) служат выражения, задающие начальную концентрацию водорода в порах материала и начальную температуру газа и порошка:

$$c(0, x) = c_0(x), \quad T(0, x) = \tilde{T}(0, x) = T_0. \quad (9)$$

В частности, можно установить  $c_0(x) = 0$ .

**3.1. Связь концентрации и давления водорода.** Для пересчёта концентрации водорода  $c(t, x)$  в давление  $p(t, x)$ , а также для определения молярного объема  $V_{\mu, 0}$  можно воспользоваться одним из уравнений состояния [8, 9], в частности, уравнением Ван-дер-Ваальса:

$$(V - nb) \left( p + \frac{an^2}{V^2} \right) = nRT. \quad (10)$$

Здесь  $n$  – количество молей газа;  $a = 0.2476 \frac{\text{bar} \cdot \text{litre}^2}{\text{mole}^2}$ ,  $b = 0.02661 \frac{\text{litre}}{\text{mole}}$  – поправки Ван-дер-Ваальса для водорода. Принято считать, что это уравнение адекватно описывает состояние водорода до давления 100–150 бар. Учитывая очевидное равенство  $cV = nN_A$  ( $c$  – концентрация в объеме  $V$ ), получаем выражение для давления водорода через концентрацию и температуру:

$$p(c, T) = \frac{RTc}{N_A - cb} - \frac{ac^2}{N_A^2}. \quad (11)$$

Это уравнение используется для определения давления водорода внутри пор материала. Давление требуется знать в случае использования более сложных моделей гидрирования, в которых ад(аб)сорбция водорода на поверхности отдельных порошинок

определяется давлением водорода. Для нахождения концентрации по давлению имеем уравнение:

$$c^3(p, T) - \frac{N_A}{b} c^2(p, T) + \frac{RT N_A^2 + bp N_A^2}{ab} c(p, T) - \frac{N_A^3 p}{ab} = 0. \quad (12)$$

Данное уравнение может быть использовано для задания начальной концентрации водорода (9) в порах порошка по известному начальному давлению. Для нахождения молярного объема по давлению имеем уравнение:

$$V_\mu^3(p, T) - \left( b + \frac{RT}{p} \right) V_\mu^2(p, T) + \frac{a}{p} V_\mu(p, T) - \frac{ab}{p} = 0. \quad (13)$$

Молярный объем водорода при данном давлении  $V_{\mu,0}$  входит в уравнение (4). Приведенные кубические уравнения из физического смысла должны иметь только один положительный вещественный корень, грубое значение которого можно искать аналитически (метод Кардано), а точное значение численно (например, методом Ньютона). В уравнениях (11)–(13) концентрация  $c$  в  $\frac{1}{litre}$  ( $1 litre = 1000 cm^3$ ), а молярный объем  $V_\mu$  в  $\frac{litre}{mole}$ .

**3.2. Учет химических реакций гидрирования и дегидрирования, вычисление пористости, плотности газа и материала.** В модели тепломассопереноса (1)–(9) неопределенной пока величиной является количество поглощенного водорода  $a(t, x)$ . В общем случае, здесь возникает отдельная подзадача моделирования поглощения водорода порошком при известном давлении, температуре, а также известных физико-химических и геометрических характеристиках порошка. Изложим простейшую модель поглощения водорода гидридом, модель дегидрирования может быть построена аналогично. Будем считать, что порошок состоит из сферических порошинок радиуса  $R_0$ . В результате гидрирования порошинки поглощают водород до некоторой максимальной величины и пропорционально поглощенному водороду меняется их объем и выделяется тепловая энергия. При этом меняется пористость порошка  $\eta(t, x)$ . Положим, что поглощение водорода гидридом происходит пропорционально текущей концентрации водорода  $c(t, x)$  с некоторым коэффициентом пропорциональности

$\Delta(t, x)$  (в простейшем случае константа) до достижения максимального содержания водорода в гидриде  $H_c$ :

$$a(t, x) = \begin{cases} c(t, x)\Delta(t, x), & t \leq t_h^*(x) \\ 0, & t > t_h^*(x). \end{cases}$$

Время окончания процесса гидрирования  $t_h^*(x)$  определяется из очевидного выражения:

$$S \int_0^\ell \int_0^{t_h^*(x)} a(\tau, x) d\tau dx = S \int_0^\ell H_c(1 - \eta(0, x)) dx.$$

Количество выделяющейся энергии при гидрировании определяется через теплоту образования гидрида  $H_E$ :

$$E_h(t, x) = a(t, x)H_E/N_A.$$

Теперь возможно вычислить плотность порошка  $\tilde{\rho}(t, x)$  и пористость  $\eta(t, x)$ . Сначала определим вспомогательные величины:

$$Q = \frac{3S\ell(1 - \eta_0)}{4\pi R_0^3}, \quad H_d(t, x) = \left( \int_0^{\min(t, t_h^*(x))} a(\tau, x) d\tau \right) \left( H_c(1 - \eta(0, x)) \right)^{-1}.$$

Здесь  $Q$  — количество всех порошинок;  $H_d(t, x)$  — доля гидрида в порошке;  $\eta_0$  — некоторая средняя начальная пористость порошка. Поскольку считаем, что объем частицы меняется пропорционально количеству поглощенного водорода, т.е. доле гидрида в порошке, то можно написать выражения для массы металла  $m_{me}$  и гидрида  $m_h$  в порошинке:

$$m_{me} = \frac{4}{3}\pi R_0^3 \tilde{\rho}_{me} (1 - H_d(t, x)), \quad m_h = \frac{4}{3}\pi R_0^3 \tilde{\rho}_{me} H_d(t, x) + \frac{\mu S \ell}{Q N_A} \int_0^{\min(t, t_h^*(x))} a(\tau, x) d\tau.$$

Отсюда возможно получить выражения для радиуса порошинки и плотности:

$$R(t, x) = \left[ \frac{R_0^3 H_d(t, x)}{\tilde{\rho}_h} \left( \tilde{\rho}_{me} + \frac{\mu H_c}{N_A} \right) + R_0^3 (1 - H_d(t, x)) \right]^{1/3},$$

$$\tilde{\rho}(t, x) = \frac{R_0}{R(t, x)} \left( \tilde{\rho}_{me} + \frac{\mu H_c H_d(t, x)}{N_A} \right)^{1/3}.$$

Далее можем определить пористость:

$$\eta(t, x) = 1 - \frac{R^3(t, x)}{R_0^3} (1 - \eta(0, x)).$$

Здесь использованы величины:  $\tilde{\rho}_{me}$ ,  $\tilde{\rho}_h$  — кристаллические плотности металла и гидрида, которые могут быть функциями температуры, если требуется учесть линейное расширение. В реальности учитывать линейное расширение смысла не имеет, поскольку вызванное им изменение плотности и пористости намного меньшее, чем при фазовых переходах во время реакций гидрирования (дегидрирования).

Наряду с такой простейшей моделью можно рассматривать более сложные модели гидрирования (дегидрирования), в которых учтены отдельные физико-химические процессы, протекающие на поверхности и в объеме порошинок, также учитывать распределение порошинок по размерам [6, 7]. В данной задаче это позволит учесть кинетику протекания реакций, однако потребует намного больших затрат времени при вычислениях.

#### 4. ИЗМЕНЕНИЕ ФИЗИЧЕСКИХ ХАРАКТЕРИСТИК МАТЕРИАЛА И ГАЗА

Поскольку температуры и концентрации меняются в широких пределах, то физические характеристики водорода и порошка меняются на десятки и сотни процентов, в модели это можно учесть следующим образом. Экспериментальные зависимости теплоемкости и теплопроводности от температуры возможно аппроксимировать рядами в виде ( $m$  — некоторое достаточно малое целое число):

$$\begin{aligned} q(t, x) &= q_0(t, x) + \sum_{i=-m}^m q_i T^i(t, x), \quad \tilde{q}_{me}(t, x) = \tilde{q}_{me,0}(t, x) + \sum_{i=-m}^m \tilde{q}_{me,i} T^i(t, x), \\ \tilde{q}_h(t, x) &= \tilde{q}_{h,0}(t, x) + \sum_{i=-m}^m \tilde{q}_{h,i} T^i(t, x), \quad \tilde{q}(t, x) = \tilde{q}_{me}(t, x)(1 - H_d(t, x)) + \tilde{q}_h H_d(t, x), \\ \lambda(t, x) &= \lambda_0(t, x) + \sum_{i=-m}^m \lambda_i T^i(t, x), \quad \tilde{\lambda}_{me}(t, x) = \tilde{\lambda}_{me,0}(t, x) + \sum_{i=-m}^m \tilde{\lambda}_{me,i} T^i(t, x), \\ \tilde{\lambda}_h(t, x) &= \tilde{\lambda}_{h,0}(t, x) + \sum_{i=-m}^m \tilde{\lambda}_{h,i} T^i(t, x), \quad \tilde{\lambda}(t, x) = \tilde{\lambda}_{me}(t, x)(1 - H_d(t, x)) + \tilde{\lambda}_h H_d(t, x). \end{aligned}$$

Возможным вариантом описания концентрационной и температурной зависимости коэффициента диффузии в некотором рабочем диапазоне концентраций и температур можно положить:

$$D(c, T) = \begin{cases} D_0(1 + \varepsilon c(t, x)) \exp(-E_D / KT(t, x)), & \eta(t, x) > \eta^*, \\ D_{me+hydr}, & \eta(t, x) \leq \eta^*. \end{cases}$$

Здесь параметры  $D_0$ ,  $E_D$ ,  $\varepsilon$  определяют зависимость для пористости выше некой критической величины  $\eta^*$ . В случае если пористость оказывается ниже, то можно считать, что порошок спекся и диффузия происходит преимущественно в материале порошка.

### 5. ЗАКЛЮЧЕНИЕ

В статье рассмотрена одномерная модель физико-химических процессов в гидридов аккумуляторе водорода. Данная модель была реализована программно. Пробные расчеты показали адекватность модели.

В данной модели использовано уравнение диффузии (дисперсии) газа в пористом материале. Альтернативным подходом к моделированию процессов переноса в пористых материалах является использование уравнений газовой динамики, т.е. рассмотрение уравнений неразрывности сжимаемого газа, движения и энергии газа [10].

### Литература

1. Лыков А.В., Михайлов Ю.А. Теория тепло и массопереноса. М-Л.: Госэнергоиздат, 1963.
2. Лыков А.В. Тепломассообмен (справочник). М.: Энергия, 1978.
3. Самарский А.А., Михайлов А.П. Математическое моделирование: Идеи. Методы. Примеры. М.: Физматлит, 2005.
4. Зигель Р., Хауэлл Д. Теплообмен излучением. М.: Мир, 1975.
5. Уонг Х. Основные формулы и данные по теплообмену для инженеров (справочник). М.: Атомиздат, 1979.
6. Заика Ю.В., Родченкова Н.И. Моделирование ТДС-спектра дегидрирования с учетом сжатия и теплопоглощения // Обозрение прикладной и промышленной математики, том 12, выпуск 4, 2005, с.957-960.
7. Заика Ю.В., Чернов И.А. Краевая задача с динамическими условиями и движущейся границей (кинетика дегидрирования) // Математическое моделирование, том 16, №4, 2004, с.3-6.
8. Вуколович М.П., Новиков И.И. Уравнение состояния реальных газов. М.-Л.:Госэнергоиздат, 1948.
9. Мейсон 7.Э., Сперлинг Е. Вириальное уравнение состояния. М.: Мир, 1972.
10. Баренблatt Г.И., Ентов В.М., Рыжик В.М. Теория нестационарной фильтрации жидкости и газа. М.: Недра, 1972.

УДК 519.68

**РЕАЛИЗАЦИЯ И ИСПОЛЬЗОВАНИЕ ЗАДАЧ  
МОДЕЛИРОВАНИЯ ВЗАИМОДЕЙСТВИЯ ВОДОРОДА  
С ТВЕРДЫМ ТЕЛОМ В СПЕЦИАЛИЗИРОВАННОМ  
ПРОГРАММНОМ КОМПЛЕКСЕ**

В. В. ПОПОВ, П. А. СЕМИН

Рассматриваются практические вопросы программной реализации и использования различных моделей взаимодействия водорода с твердым телом в специализированном программном комплексе.

**1. ВВЕДЕНИЕ**

В настоящее время в мире проводятся широкомасштабные исследования по поиску эффективных способов хранения и транспортировки водорода, который рассматривается в качестве перспективного энергоносителя в рамках концепции водородной энергетики [1–3]. Ввиду большой стоимости и трудоемкости проведения реальных экспериментов по изучению взаимодействия водорода с твердым телом частью исследовательских работ является математическое моделирование. В частности, это моделирование процессов гидрирования и дегидрирования порошков металлов в связи с конструированием гидридных накопителей (аккумуляторов) водорода, моделирование процессов проницаемости через конструкционные материалы и защитные покрытия — проблемы выбора материалов и проблемы водородной коррозии (хрупкости) металлов, решение обратных задач определения параметров моделей по экспериментальным данным. Такие работы ведутся в лаборатории МПТС ИПМИ КарНЦ РАН на протяжении последних лет [4–8]. Ввиду разнообразия и специфики задач, моделей, численных методов и алгоритмов за последние годы в лаборатории создано множество различных программ. Основным недостатком этих

программ, препятствующим дальнейшему их использованию, является "недооформленность" для использования конечным потребителем — физиком-исследователем или инженером. Возникла необходимость унифицировать и интегрировать их в единый масштабируемый программный комплекс (ПК) [9], который позволил бы эффективно хранить, редактировать, обрабатывать модельные и экспериментальные данные. В статье рассматриваются вопросы реализации в данном программном комплексе задач моделирования взаимодействия водорода с твердым телом и их использования.

## 2. Концепция программного комплекса

Исследования взаимодействия водорода с твердым телом предполагают проведение экспериментов. При этом происходящие физико-химические процессы, возможно описать с некоторой точностью в виде математической модели. В зависимости от задачи, экспериментального метода, степени детализации и учета тех или иных физико-химических процессов, математические модели могут быть различной сложности и содержать от нескольких единиц до десятков моделируемых величин (функции концентраций, десорбционные потоки и пр.), которые будут определяться множеством параметров модели (параметры, описывающие материалы и вещества, используемые в эксперименте) и проведения эксперимента. Для моделирования, идентификации параметров различных моделей и визуализации результатов могут использоваться совершенно различные алгоритмы. Таким образом, множество задаваемых параметров, особенности использования и работы тех или иных алгоритмов делают затруднительным проведение большого количества численных экспериментов и повышают вероятность ошибок совершаемых исследователем.

В качестве примера рассмотрим моделирование водородопроницаемости защитного покрытия, нанесенного на конструкционный материал (модель [10], МКИ). Для проведения расчета требуется задать:

- 12 параметров, определяющих водородопроницаемость двухслойной системы материал-покрытие (для каждого слоя задаются коэффициенты адсорбции, десорбции, диффузии, захвата и высвобождения водорода физико-химическими дефектами структуры материала и коэффициенты на стыке слоев);

- 8 параметров, задающих условия проведения эксперимента (толщины слоев, длительность эксперимента, давления в камерах, период включения диссоциатора, уровень концентрации под поверхностью при включенном диссоциаторе, начальное распределение водорода в материале);
- 12 параметров, определяющих работу алгоритма и запись результатов (количество шагов по пространству и времени, названия 2-х файлов потоков и 8-ми файлов концентраций).

Таким образом, всякий раз перед проведением расчетов требуется задать, как минимум, 32 параметра. Такое количество параметров задавать в командной строке трудоемко, поэтому лучше хранить их в файле настроек, но исследователь, проводящий расчеты, начнет пугаться после проведения десятков и сотен расчетов с использованием различных алгоритмов, требующих различных файлов настроек. Отметим, что в случае если требуется посмотреть изменение выходного десорбционного потока при изменении только одного параметра, задание остальных 31 параметра будет излишним, поскольку 10 параметров (имена файлов) можно сгенерировать автоматически, а значения 20 параметров взять из предыдущего эксперимента.

Появляется необходимость создания более удобного и более сложного пользовательского интерфейса, что требует значительных затрат сил и времени. Закономерным является сделать его универсальным и настраиваемым для различных задач, т.е. разделить программный код каждой задачи на алгоритмическую и интерфейсную части. Тогда реализация новой задачи сведется к настройке интерфейсной части и написанию кода только подключаемого алгоритма.

В процессе реализации различных задач выработалась следующая концепция программного комплекса. Программный комплекс состоит из двух частей (рис. 1):

- вычислительные проекты, каждый из которых включает в себя БД (содержит параметры моделей, экспериментов, алгоритмов и пр.), алгоритмы для расчета и визуализации (набор DLL и/или MATLAB программ), файлы настроек пользователяского интерфейса;
- ядро, включающее в себя пользовательский интерфейс (редакторы проектов, экспериментов, полей, БД), программный интерфейс (связь с БД и алгоритмами оформленными в виде DLL или MATLAB программ), а также общую функциональность для всех вычислительных задач.

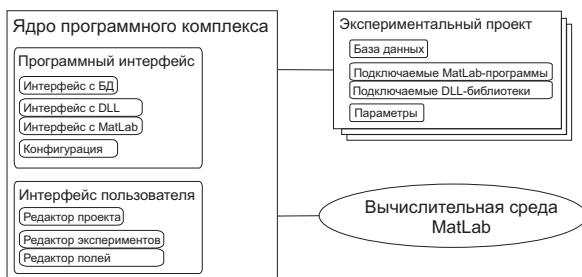


Рис. 1. Структура программного комплекса.

Пользовательская работа с программным комплексом выглядит следующим образом. Пользователь запускает ПК, открывает в нем нужный вычислительный проект, при этом подключается БД, настраивается интерфейс. Теперь в программном комплексе возможно проводить расчеты, предусмотренные выбранным проектом (рис. 5). Далее пользователь должен создать и заполнить записи в БД. Каждая запись соответствует вычислительному эксперименту и хранит все необходимые параметры. При этом часть полей может заполняться автоматически генерируемыми значениями или значениями по умолчанию. При работе с БД пользователь может настраивать и использовать сервисные функции (копирование и вставка значений групп полей, заполнение по умолчанию, автоматическая генерация значений, сортировка, поиск, группировки, отборы). После задания параметров остается запустить алгоритм для расчета и просмотреть полученные результаты нажатием соответствующих кнопок.

Для разработчика реализация алгоритма сводится к созданию вычислительного проекта. С помощью встроенных средств программного комплекса он создает БД и настраивает пользовательский интерфейс. Далее следует реализация алгоритма в виде DLL или EXE файла на любом языке программирования или написание MATLAB программы. При этом разработчик должен ориентироваться на работу DLL с программным комплексом через программный интерфейс, который позволяет управлять работой алгоритма, реализует передачу и возврат параметров (передается вся выбранная запись, настройки проекта и ПК). Для использования средств визуализации предусмотренных в ПК, следует формировать файлы данных некоторого предусмотренного формата. В случае написания алгоритмов в виде DLL

(EXE) в среде Delphi или в виде MATLAB программ можно использовать разработанные классы и библиотеки функций, что облегчает использование программного интерфейса и работу с файлами.

Такой подход позволяет значительно уменьшить количество рутинной работы при реализации и использовании алгоритмов. Пользователь имеет удобный единообразный для всех задач графический интерфейс с богатой настраиваемой функциональностью. Работа пользователя состоит из последовательности действий одинаковых для любых расчетов. Отпадает необходимость помнить настройки отдельных алгоритмов и специфику задач. Параметры и результаты расчетов удобно хранятся, имеется возможность ориентации в проведенных вычислительных экспериментах (поиск, сортировка, отборы, группировки параметров). Разработчик сосредотачивается только на разработке вычислительного алгоритма и таким образом сокращаются затраты времени и сил. Отпадает необходимость в перекомпиляции всего программного комплекса и есть возможность скрыть детали реализации алгоритмов. Для реализации алгоритмов возможно использование любых языков программирования.

### 3. ОСНОВНЫЕ ЭТАПЫ РЕАЛИЗАЦИИ АЛГОРИТМА

Для реализации алгоритма в программном комплексе требуется создать вычислительный (экспериментальный) проект, который будет включать в себя БД и файлы настроек, после чего создать DLL (EXE) или MATLAB программу. Рассмотрим последовательно все этапы.

#### **Создание экспериментального проекта.**

Создание экспериментального проекта начинается с выбора пункта меню **File\New**. Далее в появившемся диалоговом окне выбора файла проекта (рис. 2) требуется указать каталог размещения проекта и названия файла. Далее откроется окно настройки проекта (рис. 2), в котором следует заполнить все поля. Если какие-либо указанные файлы не существуют, то они будут автоматически созданы. При этом автоматически создается пустая база данных, которую возможно редактировать во встроенным редакторе (рис. 3). Требуется задать и настроить все необходимые поля. Названия полей будут являться названиями параметров в списках параметров передаваемых в код алгоритма. Далее необходимо указать пути к алгоритмам и к файлам данных на закладке **Algorithms\Parameters** (рис. 3). Настройка на

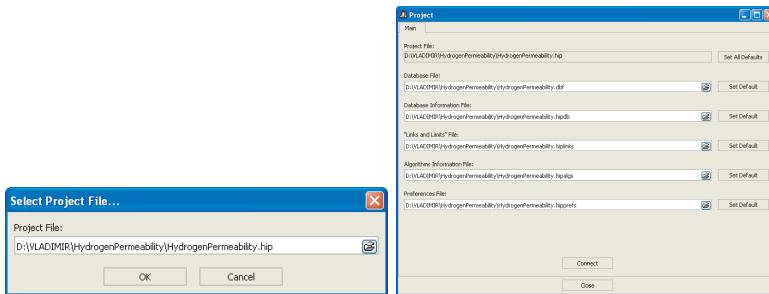


Рис. 2. Окна выбора и настройки проекта.

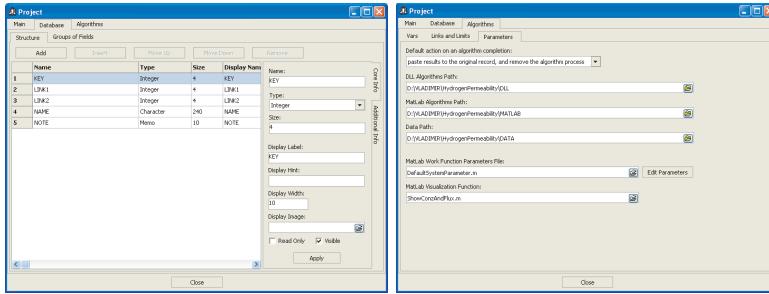


Рис. 3. Окна редактора БД и настройки дополнительных параметров проекта.

других закладках групп полей, переменных проекта, связей и ограничений полей является необязательным, но полезным.

После выхода из диалогового окна настройки проекта в указанном ранее каталоге проекта будут сформированы следующие файлы:

- основной файл проекта (расширение .hip), в котором хранятся ссылки на другие файлы и некоторые параметры проекта;
- файлы базы данных (.dbf,.cdx,.fpt), в которых хранятся данные численных экспериментов, индексы и комментарии;
- файл конфигурации полей БД (.hipdb), который содержит описание полей БД;
- файл связей и ограничений (.hiplinks), в котором определяются связи между полями БД (и их значениями), а также ограничения значений для полей БД;
- файл описания алгоритмов (.hipalgs), определяющий настройки алгоритмов и некоторые другие параметры;

- файл вспомогательных настроек (.hipref).

Теперь настройка проекта закончена, при необходимости можно открыть проект (**File\Open**, выбрать .hip файл) и поменять любые параметры или изменить БД проекта (**File\Project\Settings**).

### **Вызов алгоритмов из ПК и передача параметров.**

Программный комплекс может использовать алгоритмы, оформленные как исполняемые файлы, динамически подключаемые библиотеки (DLL) и MATLAB программы. Комплекс определяет реализацию алгоритма по расширению файла (.exe,.dll,.m) и в зависимости от этого варьируется способ вызова алгоритма и обмена данными. В общем случае всегда может использоваться обмен данными через файлы определенного формата. Это позволило упростить и сделать гибче код, устраниТЬ потенциальную проблему несовместимости со средой MATLAB других версий, облегчить отладку отдельных алгоритмов.

При вызове алгоритма программный комплекс выполняет следующие действия:

- создает следующие уникальные временные файлы:
  - файл с входными данными алгоритма;
  - файл индикатора работы алгоритма;
  - файл для сохранения результатов работы алгоритма;
- запускает выполнение алгоритма в отдельном процессе, что позволяет выполнять параллельно несколько алгоритмов;
- передает алгоритму имена созданных файлов, а также некоторые другие параметры;
- ожидает завершения работы алгоритма. Работа алгоритма может быть прервана по желанию пользователя;
- в случае успешного завершения, считывает и сохраняет результаты работы алгоритма, удаляет временные файлы.

Запуск алгоритмов, реализованных в виде исполняемых файлов, производится с помощью функции API CreateProcess(). Название файла с параметрами передается в командной строке.

Выполнение алгоритмов, оформленных в виде DLL, инициируется вызовом экспортirуемой функции TDLLMainProc строго определенного формата. Все необходимые параметры передаются через аргументы этой функции.

Для выполнения MATLAB программ комплекс использует модуль MATLAB-launcher, оформленный в виде DLL, который через MATLAB API запускает вычислительную среду, вызывает функцию

MatLabAlgorithmLauncher и передает ей названия файла параметров. Функция считывает файл параметров, прописывает пути в MATLAB и запускает указанную в файле функцию с алгоритмом (рис. 4).

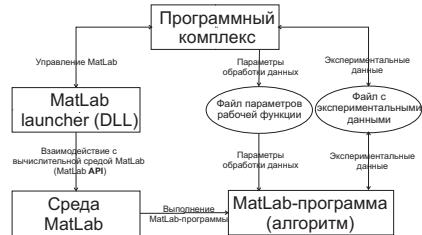


Рис. 4. Запуск MATLAB программы.

Файл параметров, который используется с MATLAB программами и исполняемыми файлами, состоит из строк вида <имя параметра> - <значение параметра>. Обязательным является наличие параметра с именем Function name, который задает MATLAB функцию с алгоритмом. Кроме этого в файле могут быть переданы параметры, необходимые для вызываемой функции (при запуске вычислительного алгоритма передаются названия полей и их значения для выбранной записи БД, а при визуализации результатов — название файла данных), и параметры с именами Path.<номер> для указания путей, по которым MATLAB будет искать вспомогательные функции, файлы параметров и данных. При завершении работы алгоритм может переписать этот файл параметров, записав в него в том же формате новые значения параметров. Программный комплекс считывает параметры и записывает их в выбранную запись БД. Если алгоритм в файле параметров не находит нужный параметр, то произойдет ошибка, которую обрабатывает алгоритм. Программный комплекс игнорирует ошибки, возникающие при считывании возвращаемых параметров.

Поскольку система MATLAB имеет мощные средства визуализации, то этот механизм вызова MATLAB используется в программном комплексе для просмотра результатов работы. При этом в качестве названия функции с алгоритмом передается название функции для построения графиков, а в качестве параметров имя файла данных.

### **Создание алгоритма в виде DLL или EXE файла.**

Для определенности будем рассматривать написание алгоритма в среде Delphi 7. Создание DLL в Delphi начинается с пункта меню (**File\New\DLL**), после чего требуется написать основной модуль:

```

unit AlgDLL;

interface uses MyDefsFile, Classes, SysUtils; //указываем свои
файлы с кодом, а также
                                // необходимые стандартные библиотеки
procedure TDLLMainProc(var Options: APOptions; Vars: TList;
Params: TList;
  var fAbort: boolean; var ExecStatus: integer; ExecMessage: PChar); StdCall;
exports TDLLMainProc;

implementation procedure TDLLMainProc(var Options: APOptions;
Vars: TList; Params: TList;
  var fAbort: boolean; var ExecStatus: integer; ExecMessage: PChar); StdCall;
var Alg: TMyAlgorithm; Begin
  if (ExecStatus <= 0)and(not fAbort) then
  Begin
    Try //Создаем экземпляр класса и инициализируем переменные
      Alg:=TMyAlgorithm.Create(Options, Vars, Params);
    Try
      Alg.Computation;                                //Вычисляем
    except
      Alg.ErrorStatus := 1;
      Alg.ErrorMessage := 'Неизвестное исключение';
    end;
    ExecStatus := Alg.ErrorStatus;
    StrCopy(ExecMessage, PChar(Alg.ErrorMessage));
    Params := Alg.GetParams(Params); //Дописываем в список Params,
  finally
    Alg.Free;                                         //возвращаемые значения
    //Уничтожаем экземпляр класса
  end;
end;
end; end.
```

Здесь описывается основная экспортруемая из DLL процедура TDLLMainProc, которая в качестве параметров имеет структуру Options, списки параметров Vars и Params — переменные проекта и поля выбранной записи в БД, флаг прерывания работы fAbort (не используется), код и сообщение об ошибке: ExecStatus, ExecMessage. Описание класса TMyAlgorithm, реализующего алгоритм, а также структур APOptions и APVal, являющимися пунктами списков TList, находятся в файле MyDefsFile. Ниже приводим описание структур, в которых передаются данные между программным комплексом и DLL.

```

APVal = record //аналог struct в C++
  Name: AnsiString; //название параметра или поля БД
  Value: Variant;   //значение параметра или поля БД в выбранной записи
```

```

ValueType: TFieldType; //тип поля БД
end;

APOptions = record
  SemaphoreName: AnsiString; //имя семафора для прерывания алгоритма
  ParamsFile: AnsiString; //файл параметров
  ProgressFile: AnsiString; //файл индикатора работы
 LogFile: AnsiString; //файл сообщений, возникающих при работе
  ResultFile: AnsiString; //файл с возвращаемыми результатами работы
end;

```

Код алгоритма находится в классе TMyAlgorithm, реализация которого определяется исключительно разработчиком. Необходимым является только наличие методов: Create — для создания экземпляра класса и инициализации переменных, часть значений копируется из переданных структур; Computation — для выполнения расчетов; GetParams — для возвращения предусмотренных значений в программный комплекс, где они будут записаны в соответствующие поля выбранной записи БД; Free — для корректного уничтожения экземпляра класса. Также класс должен иметь переменные ErrorStatus и ErrorMessage для возвращения кода и сообщения об ошибке, возникшей при работе. Для завершения алгоритма по нажатию кнопки остановки из программного комплекса функция Computation должна содержать фрагмент кода, прерывающего алгоритм (переменная hndl объявлена как THandle):

```

//инициализация семафора
hndl:=OpenSemaphore(MUTANT_ALL_ACCESS, false, PChar(Options.SemaphoreName));
if hndl=0 then
hndl:=CreateSemaphore(nil,1,1,PChar(Options.SemaphoreName));
//основной цикл, в котором происходят вычисления по разностной схеме
for j := 0 to StT do
  //код разностной схемы
  ...
  //опрос семафора для остановки расчетов
  if j mod InquiryStep = 0 then //позволяет считать быстрее
    if WaitForSingleObject(hndl,1) = WAIT_TIMEOUT then
      Begin
        ErrorStatus := 10;
        ErrorMessage := 'Выполнение прервано пользователем.';
      end
    else ReleaseSemaphore(hndl,1,nil);
  //Выйдем, если произошла ошибка или прерывание пользователем
  if ErrorStatus > 0 then break;
end;

```

Создание исполняемого файла с алгоритмом производится аналогичным образом. При этом не объявляется функция TDLLMainProc. Для передачи параметров из программного комплекса считывается и разбирается файл параметров, название которого передается в командной строке.

Информировать программный комплекс о текущем состоянии вычислений, о возникающих ошибках и сообщениях алгоритмы могут через файлы, имена которых содержатся в полях ProgressFile,LogFile структуры APOptions.

### **Создание алгоритма в виде MATLAB программы.**

При запуске MATLAB программы сначала будет запущена функция MatLabAlgorithmLauncher, которая после считывания файла параметров и установки необходимых путей запустит указанную ей функцию с вычислительным алгоритмом. Код функции следующий:

```
function Status = MatLabAlgorithmLauncher(varargin) strName = '';
if (nargin == 1) strName = varargin{1}; end; if strName == ''
strName = 'Launch.par'; end; [Status, LaunchPar] =
ReadLaunchParams(strName);
% прописываем пути к функциям и файлам данных в MATLAB
for i = 1:length(LaunchPar)
    if length(findstr(LaunchPar(i).Name,'Path.')) > 0
        addpath(LaunchPar(i).Val);
    end
end; if Status == 0
    f = fcncchk(GetParamValByName('Function name', LaunchPar));
    Status = feval(f, LaunchPar);
end;

function [Status, LaunchPar] = ReadLaunchParams(FName) Status =
0; mes = ''; [F, mes] = fopen(FName, 'rt'); if F ~= -1
    % Если файл открыт, то пытаемся считать заголовок
    Str = ''; Str1 = '';
    while (feof(F)) == 0
        lines = fgetl(F); line = strrep(lines,(','),('.'));
        if findstr(' ', line) > 0  Str = char(Str, line); end
    end
    fclose(F);
    % Преобразуем заголовочную информацию в другой вид
    sz = size(Str);
    for i = 1:sz(1)
        p = findstr(' ',Str(i,:)); s = length(Str(i,:));
        Name = deblank(Str(i,1:p-1)); V1 = deblank(Str(i,p+2:s));
        LaunchPar(i).Name = char(Name); LaunchPar(i).Val = char(V1);
```

```

    end
else
    % Покажем сообщение об ошибке
    disp(mes); Status = 1;
end;

function outParVal = GetParamValByName(Name, Par) outParVal = '';
f1 = 0; Names = fieldnames(Par); for i=1:size(Par,2) if
strcmp(lower(deblank(getfield(Par(i),char(Names(1))))),lower(deblank(Name)))==1
    outParVal = getfield(Par(i),char(Names(2))); f1 =1; break; end
end if f1 == 0 disp(['параметр ' Name ' не найден!']); end;

```

Функция с вычислительным алгоритмом должна в качестве аргумента получать список с параметрами LaunchPar и возвращать код ошибки или 0 в случае нормального завершения. Получение значений параметров по имени из списка LaunchPar может осуществляться использованием функции GetParamValByName.

Информировать программный комплекс о текущем состоянии вычислений, о возникающих ошибках и сообщениях алгоритмы могут через файлы, имена которых передаются в параметрах с именами Progress file, Log file.

### **Файлы данных и их визуализация.**

Разработчик проекта может предусмотреть использование файлов данных любого формата, в том числе и своего собственного, в этом случае он должен также разработать программу для визуализации данных. Для избежания этих рутинных действий возможно использовать файлы данных предусмотренного ПК формата.

Файл данных записывается в формате ASCII и состоит из текстового заголовка, описывающего параметры, при которых файл был получен, за которым следует часть с данными. В первой строке заголовка указывается тип данных (десорбционный поток, концентрация в ловушках и т.д.), далее разделенные пустыми строками располагаются секции: Parameters, Variables, Additional parameters, Data part description. Каждая из секций включает параметры, записанные в виде <имя параметра> - <значение параметра>. Секция Parameters содержит значения полей выбранной записи БД. В секцию Variables помещаются переменные проекта. Секция Additional parameters зарезервирована. Секция Data part description содержит параметры, используемые при визуализации (значения по осям, названия осей и пр.). Тег <Data part> отделяет заголовок от части с данными. Данные записываются в виде одномерного или двухмерного массива. В случае

одномерного массива каждое число записывается с новой строки, в двухмерном массиве числа в строке разделяются ";".

При реализации алгоритмов на Delphi код для создания и записи файлов такого формата может быть следующим.

```
unit MyDefsFile;
...
const
  FT_c = 'Hydrogen concentration file';
Type
  TMyAlgorithm = class
    AddParams: TList; //Список дополнительных параметров
    Conz: array of double; //Массив, хранящий концентрацию при расчетах
    ...
  public
    f_c: TDataFile; //Объявление нашего класса, работающего с файлом
    ...
    constructor Create(var Options: APOptions; Vars: TList; Params: TList);
    destructor Destroy();
  end;
  ...
constructor TMyAlgorithm.Create(var Options: APOptions; Vars: TList; Params: TList);
var ConzFN: string;
begin
  ...
  ConzFN := GetValByName('CONZFILE', Params); //Считывание имени файла данных
  //Создание экземпляра класса
  f_c := TDataFile.Create(Params, Vars, AddParams, ConzFN, FT_c, 0);
  f_c.Clear; //Инициализация, используется перед любым открытием файла
  f_c.OpenFileToWrite; //Открытие файла, здесь пишется заголовок
  ...
end;

destructor TMyAlgorithm.Destroy;
begin
  f_c.CloseDataFile; //Закрытие файла
  f_c.Free; //Уничтожение экземпляра класса
  ...
end;

procedure TMyAlgorithm.Computation;
begin
  //основной цикл, в котором происходят вычисления по разностной схеме
  for j := 0 to StT do
  ...
  //запись данных в файл
```

```
    if j mod WriteStep = 0 then f_c.WriteNext(Conz);
end;
end;
end.
```

Построение двух и трехмерных графиков для таких файлов данных возможно с помощью MATLAB функции ShowDatFile (рис. 6).

#### 4. ИСПОЛЬЗОВАНИЕ ПРОЕКТА

При работе с комплексом исследователь использует пользовательский интерфейс (рис. 5), основной частью которого являются редакторы вычислительных экспериментов и полей. Поскольку оптимальным явилось использование БД в виде одной большой таблицы, где строки соответствуют вычислительным экспериментам, а столбцы задаваемым параметрам, то редактор экспериментов позволяет просматривать множество записей БД, но при этом видна только часть полей, а для просмотра всех полей текущей (выбранной) записи используется редактор полей. С помощью панели управления БД и любого из этих редакторов возможно редактировать данные.

Управлять данными, вычислениями и просмотром результатов пользователь может с помощью пунктов меню или с помощью кнопок на панелях инструментов. Основные действия пользователя при использовании проекта были приведены в пункте 2. В окне настройки проектов можно дополнительно задать группы копирования и вставки полей, а также группы отображения полей в редакторе полей.

#### 5. РЕАЛИЗАЦИЯ

Программный комплекс реализован в среде программирования Borland C++Builder 6.0 с использованием дополнительных компонентов: а) RXLib 2.75 для реализации специализированных типов полей; б) TMS Component Pack 3.7 для реализации интерфейса пользователя; в) Advantage Database Server 8.0 для взаимодействия с таблицами БД. Программная подсистема MATLAB-launcher оформлена в виде DLL, которая написана на языке MS Visual C++. Таблицы БД экспериментальных проектов хранятся в формате DBF/CDX.

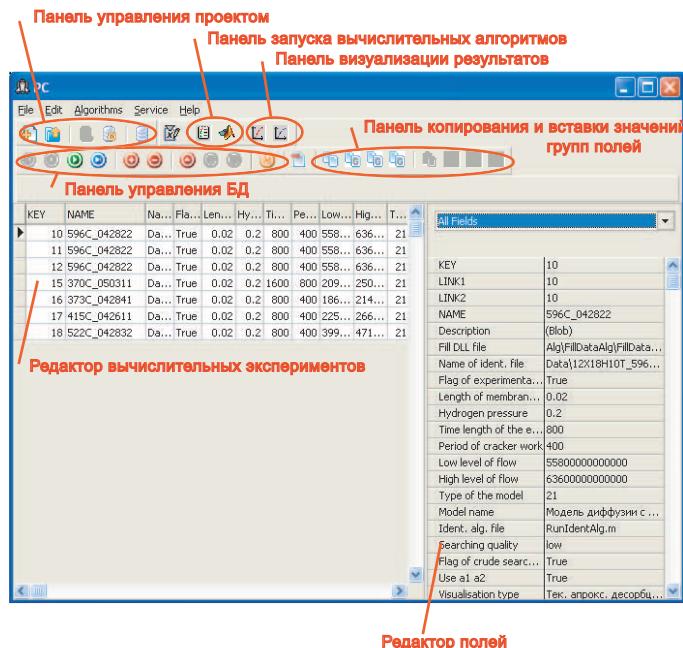


Рис. 5. Основное окно программного комплекса, редакторы экспериментов и полей.

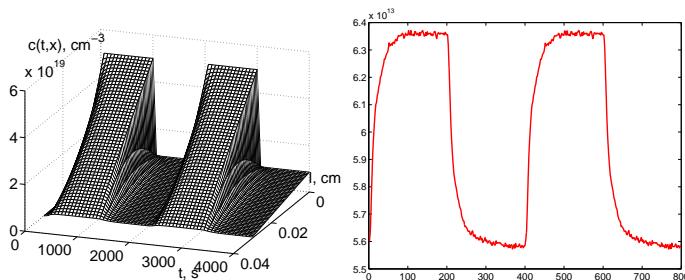


Рис. 6. Визуализация данных средствами MATLAB: объемная концентрация и десорбционный поток.

## Литература

1. International Symposium on Metal Hydrogen Systems, Fundamentals and Applications. Book of abstracts // Krakow, Poland, 2004.
2. Hydrogen Materials Science and Chemistry of Carbon Nanomaterials // NATO Science Series II: Mathematics, Physics and Chemistry. Vol. 172. Kluwer Academic Publishers, Netherlands. 2004.
3. Hydrogen Materials Science and Chemistry of Carbon Nanomaterials. Сборник тезисов IX международной конференции // Ред. Щур Д.В., Загинайченко С.Ю., Везироглу Т.Н. Kiev: AHEU, 2005.
4. Zaika Yury V. Identification of a hydrogen transfer model with dynamical boundary conditions // International Journal of Mathematics and Mathematical Sciences. 2004. N 4. Pp. 195-216.
5. Заика Ю.В. Определение параметров переноса водорода сквозь мембранны методом концентрационных импульсов // Известия вузов. Физика. 2002. N 1. С. 81-87.
6. Заика Ю.В. Параметрическая регуляризация модели водородопроницаемости с динамическими граничными условиями // Математическое моделирование. 2001. Т. 13. N 11. С. 69-87.
7. Попов В.В. Параметрическая идентификация распределенных моделей водородопроницаемости. Диссертация на соискание ученой степени кандидата физ.-мат. наук. Петрозаводск: ПетрГУ, 2004.
8. Чернов И.А. Математическое моделирование взаимодействия водорода с твердым телом (термодесорбционная спектрометрия). Диссертация на соискание ученой степени кандидата физ.-мат. наук. СПб: СПбГУ, 2004.
9. В. В. Попов, П. А. Семин Программный комплекс моделирования взаимодействия водорода с твердым телом // Труды ИПМИ КарНЦ РАН. Петрозаводск, 2005. Вып.6. С.58-68.
10. В. В. Попов Идентификация параметров моделей водородопроницаемости двухслойных мембран // Труды ИПМИ КарНЦ РАН. Петрозаводск, 2005. Вып.6. С.69-77.

УДК 519.2

**О ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИЦИРУЕМОСТИ  
МОДЕЛИ ФАКТОРНОГО АНАЛИЗА  
С ЗАВИСИМЫМИ ФАКТОРАМИ И ОСТАТКАМИ**

С. В. СТАФЕЕВ

Рассматривается модель факторного анализа с зависимыми факторами, в которой зависимость между остатками описывается ковариационной графовой моделью. Получены условия локальной идентифицируемости модели и найден класс моделей, удовлетворяющих полученным условиям идентифицируемости.

**1. ВВЕДЕНИЕ**

В работах [1, 2] рассматривались модели факторного анализа с независимыми факторами, в которых зависимость между остатками описывается гауссовской наследственной графовой моделью [3]. В данной статье мы будем рассматривать модель с зависимыми факторами, в которой зависимость между остатками представляется в виде гауссовской ковариационной графовой модели. Модель является обобщением модели, которая была рассмотрена в работе [4]. Для рассматриваемой модели будут получены достаточные условия локальной параметрической идентифицируемости модели и найден класс моделей, удовлетворяющих полученным условиям. Данный класс составляют леса при некоторых ограничениях на число вершин смежных с данной вершиной. В работе используются обозначения, введенные в работе [2].

## 2. ОПИСАНИЕ КЛАССА РАССМАТРИВАЕМЫХ МОДЕЛЕЙ

Пусть  $\widehat{\mathbf{G}}_{\mathbf{Y}}^{\mathbf{H}}$  множество всех наследственных графов [2,3] с  $n+k$  вершинами, каждый граф  $G = (\mathbf{X}, \mathbf{E})$  которого обладает следующими свойствами:

- (1)  $\mathbf{X} = \mathbf{H} \cup \mathbf{Y}$ , где  $\mathbf{H} = \{H_1, \dots, H_k\}$ , а  $\mathbf{Y} = \{Y_1, \dots, Y_n\}$ ;
- (2) Вершины графа  $G$ , которые соединены неориентированными ребрами, составляют множество  $\mathbf{H}$ ;
- (3) Любые две различные вершины множества  $\mathbf{H}$  соединены ребром;
- (4) Для любой вершины  $H' \in \mathbf{H}$ ,  $\text{Ch}(H') \neq \emptyset$ ;
- (5) Каждая вершина  $Y \in \mathbf{Y}$  имеет только одного родителя, причем

$$\mathbf{Pa}(Y) = \{H''\}, \quad H'' \in \mathbf{H}.$$

Пусть  $\{G, N_G(m, \Lambda, B, \Omega)\}$ ,  $G = (\mathbf{X}, \mathbf{E}) \in \widehat{\mathbf{G}}_{\mathbf{Y}}^{\mathbf{H}}$ ,  $\mathbf{X} = \mathbf{H} \cup \mathbf{Y}$ ,  $m \in \mathbb{R}^{n+k}$ ,  $\Sigma_G(\Lambda, B, \Omega) \in \mathbf{P}_G$ , – наследственная гауссовская графовая модель [2,3] с множеством латентных случайных величин  $\mathbf{H}$  и множеством наблюдаемых случайных величин  $\mathbf{Y}$ .

Удалим в графе  $G$  все вершины множества  $\mathbf{H}$  и инцидентные с ними ребра. Полученный двуориентированный граф  $S = (\mathbf{Y}, \mathbf{E}_S)$  мы будем называть структурой модели.

Мы можем представить зависимость между элементами множества  $\mathbf{Y}$  в виде системы регрессионных уравнений

$$Y_i = m_i + a_i H(Y_i) + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

где

$$H(Y_i) = \mathbf{Pa}(Y_i);$$

$m = \{m_1, \dots, m_n\}$  – вектор математических ожиданий наблюдаемых случайных величин.

Зависимость между компонентами случайного вектора остатков  $\varepsilon = \{\varepsilon_1, \dots, \varepsilon_n\}$  описывается гауссовской ковариационной графовой моделью со структурой  $S$ . Пусть  $\Omega = (\omega_{ij})$  – матрица ковариаций вектора  $\varepsilon$ , тогда

$$\text{если } Y_i \leftrightarrow Y_j \notin \mathbf{E}, \quad \text{то } \text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad (\omega_{ij} = 0).$$

Случайные величины  $H_1, \dots, H_k$  имеют общее нормальное распределение с матрицей ковариаций  $A = (a_{ij})$ , причем  $\mathbf{M}(H_i) = 0$  и  $\mathbf{Var}(H_i) = 1$ ,  $i = 1, \dots, k$ .

Пусть  $\theta = \{m, a_i, i = 1, \dots, n, \Omega, A\}$ . Определим параметрическое множество  $\Theta = \{\theta | \Omega, A - положительно определены.\}$

Мы будем говорить, что модель (1) идентифицируема при  $\theta \in \Theta' \subset \Theta$ , если при  $\theta \in \Theta' \subset \Theta$  по матрице ковариаций наблюдаемых случайных величин  $\Sigma_{\mathbf{Y}}$  параметры  $\{\omega_{ij}, Y_i \leftrightarrow Y_j \in \mathbf{E}_S\}$  определяются однозначно, а параметры  $\{a_1, \dots, a_n, a_{ij}, 1 \leq i < j \leq k\}$  с точностью до знака.

Мы будем говорить, что модель (1) локально идентифицируема при  $\theta \in \Theta'' \subset \Theta$ , если для любого  $\theta \in \Theta''$  модель будет идентифицируемой при  $\theta \in U_\theta$ , где  $U_\theta$  – некоторая окрестность  $\theta$ .

В работе [4] были получены условия идентифицируемости модели (1) для случая  $k = 1$ . Мы будем рассматривать случай произвольного  $k$ .

Определим  $k$  множеств

$$\mathbf{Y}^i = \{Y \in \mathbf{Y} | H_i = \mathbf{Pa}(Y)\}$$

и пусть  $|\mathbf{Y}^i| = n_i$ ,  $i = 1, \dots, k$ .

Пусть  $\overline{S}(\mathbf{Y}, \overline{\mathbf{E}})$  – дополнительный граф структуры  $S$ .

Пусть  $C$  подграф дополнительного графа, являющийся простым циклом, состоящим из вершин  $\{V_1, \dots, V_k\} \in \mathbf{Y}$ , а  $G$  подграф дополнительного графа, состоящий из простого цикла с вершинами  $\{V_1, \dots, V_{i_1}\} \in \mathbf{Y}$ , простой цепи с вершинами  $\{V_{i_1}, \dots, V_{i_2}\} \in \mathbf{Y}$  и простого цикла с вершинами  $\{V_{i_2}, \dots, V_n\} \in \mathbf{Y}$ .

Для  $1 \leq m < l \leq k$  определим функции

$$F_1^{ml}(C) = \sum_{i=1}^{k-1} \chi_{ml}(V_i, V_{i+1})(-1)^{i+1} + \chi_{ml}(V_1, V_k)(-1)^{k+1}$$

и

$$F_2^{ml}(G) = \begin{aligned} & \sum_{j=1}^{i_1-1} \chi_{ml}(V_j, V_{j+1})(-1)^{j+1} + \chi_{ml}(V_1, V_{i_1})(-1)^{i_1+1} + \\ & 2 \left( \sum_{j=i_1}^{i_2-1} \chi_{ml}(V_j, V_{j+1})(-1)^{j+1} \right) + \\ & \sum_{j=i_2}^{n-1} \chi_{ml}(V_j, V_{j+1})(-1)^{j+1} + \chi_{ml}(V_n, V_{i_2})(-1)^{i_2}, \end{aligned}$$

где  $\chi_{ml}(V_i, V_j) = 1$  если  $V_i \in \mathbf{Y}^m$ ,  $V_j \in \mathbf{Y}^l$  или  $V_j \in \mathbf{Y}^m$ ,  $V_i \in \mathbf{Y}^l$  и  $\chi_{ml}(V_i, V_j) = 0$  в противном случае.

### 3. УСЛОВИЯ ИДЕНТИФИЦИРУЕМОСТИ ПРИ $k = 2$

Пусть множество  $\mathbf{H}$  состоит из двух вершин (т.е.  $k = 2$ ) и пусть  $a_{ij} = a$  и  $F_1^{12}(C) = F_1(C)$ ,  $F_2^{12}(G) = F_2(G)$ . Для этого случая нами получены необходимые и достаточные условия идентифицируемости модели (1).

**Теорема 1.** Пусть  $k = 2$ ,  $a_i \neq 0$ ,  $a \neq 0$ ,  $i = 1, \dots, n$ . Модель 1 локально идентифицируема тогда и только тогда, когда

- (1) каждая компонента связности дополнительного графа  $\overline{S}$  содержит по крайней мере один нечетный цикл;
- (2) по крайней мере одна компонента связности содержит либо четный цикл  $C$ , для которого  $F_1(C) \neq 0$ , или подграф  $G$ , состоящий из двух простых нечетных циклов, соединенных простой цепью, для которого  $F_2(G) \neq 0$ .

**Доказательство.** Мы будем использовать регрессионное представление модели (1). Пусть  $\Sigma = (\sigma_{ij})$  – матрица ковариаций наблюдаемых случайных величин. Используя (1), мы получаем

$$\sigma_{ij} = \begin{cases} \alpha_{ij}a_i a_j + \omega_{ij}, & \text{если } Y_i - Y_j \notin \overline{\mathbf{E}}, \\ \alpha_{ij}a_i a_j, & \text{если } Y_i - Y_j \in \overline{\mathbf{E}}, \end{cases} \quad (2)$$

где

$$\alpha_{ij} = \begin{cases} 1, & \text{если } Y_i, Y_j \in \mathbf{Y}^1 \text{ или } Y_i, Y_j \in \mathbf{Y}^2, \\ a, & \text{в противном случае.} \end{cases}$$

Система (2) задает отображение

$$\varphi : \Theta \rightarrow \mathbb{R}^{\frac{n(n+1)}{2}}.$$

Матрица Якоби  $J$ , соответствующая отображению  $\varphi$ , имеет следующий блочно-верхне-диагональный вид

$$J = \begin{pmatrix} J_1 & * \\ & J(\alpha) \end{pmatrix},$$

где

$$J_1 = \left( \frac{\partial \sigma_{ij}}{\partial \omega_{ij}} \right)_{Y_i - Y_j \notin \overline{\mathbf{E}}} = I_{\frac{n(n+1)}{2} - |\overline{\mathbf{E}}|},$$

$$J(\alpha) = \left( \frac{\partial \sigma_{ij}}{\partial a_1}, \dots, \frac{\partial \sigma_{ij}}{\partial a_n}, \frac{\partial \sigma_{ij}}{\partial a} \right)_{Y_i - Y_j \in \overline{\mathbf{E}}}.$$

Модель (1) будет локально идентифицируемой тогда и только тогда, когда матрица  $J$  имеет полный ранг [5]. (Под полным рангом мы подразумеваем полный ранг по столбцам, т.е. когда ранг матрицы равен числу столбцов матрицы.) Легко видеть, что в нашем случае это эквивалентно тому, что матрица  $J(\alpha)$  имеет полный ранг.

Легко видеть, что для  $i, j$ ,  $Y_i - Y_j \in \overline{\mathbf{E}}$ ,  $k = 1, \dots, n$ ,

$$\frac{\partial \sigma_{ij}}{\partial a_k} a_k = d_{ij}^k a_{ij} a_i a_j, \quad \frac{\partial \sigma_{ij}}{\partial a} a_{ij} = h_{ij} a_{ij} a_i a_j,$$

где для  $k = 1, \dots, n$

$$d_{ij}^k = \begin{cases} 1, & \text{если } k = i, j, \\ 0, & \text{если } k \neq i, j, \end{cases}$$

и

$$h_{ij} = \begin{cases} 1, & \text{если } a_{ij} = a, \\ 0, & \text{если } a_{ij} = 1. \end{cases}$$

Пусть  $d_{ij} = (d_{ij}^1, \dots, d_{ij}^n)$  и  $w_{ij} = (d_{ij}, h_{ij})$ .

Образуем две матрицы  $D = (d_{ij})_{Y_i - Y_j \in \overline{\mathbf{E}}}$  и  $W = (w_{ij})_{Y_i - Y_j \in \overline{\mathbf{E}}}$ . Очевидно, что

$$J(\alpha) \operatorname{diag}\{a_1, \dots, a_n, a\} = \operatorname{diag}\{a_{ij} a_i a_j, Y_i - Y_j \in \overline{\mathbf{E}}\} W.$$

Ввиду того, что  $a_i \neq 0$ ,  $i = 1, \dots, n$ ,  $a \neq 0$ , матрица  $J(\alpha)$  имеет полный ранг тогда и только тогда, когда матрица  $W$  имеет полный ранг.

Заметим, что используя (2), можно получить

$$(\ln |\sigma_{ij}|)_{Y_i - Y_j \in \overline{\mathbf{E}}} = W \begin{pmatrix} \ln |a_1| \\ \vdots \\ \ln |a_n| \\ \ln |a| \end{pmatrix}.$$

Откуда получаем, что если матрица  $W$  имеет полный ранг, то параметры  $\{a_1, \dots, a_n, a\}$  определены с точностью до знака.

Пусть дополнительный граф  $\overline{S}$  состоит из  $l$  компонент связности  $\overline{S}_1 = (\mathbf{Y}^1, \overline{\mathbf{E}}_1), \dots, \overline{S}_l = (\mathbf{Y}^l, \overline{\mathbf{E}}_l)$ . Легко видеть, что матрица  $D$

имеет блочно-диагональный вид

$$D = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_l \end{pmatrix}$$

где  $D_i$ ,  $i = 1, \dots, l$ , матрицы, соответствующие компонентам связности графа  $\bar{S}$ .

Матрица  $D$  имеет полный ранг тогда и только тогда, когда каждая из матриц  $D_1, \dots, D_l$  имеет полный ранг.

Пусть  $\bar{S}'_i$  связный подграф графа  $\bar{S}_i$  с  $\mathbf{Y}^i$  вершинами и содержащий  $|\mathbf{Y}^i|+1$  ребер и пусть  $\bar{S}''_i$  подграф  $\bar{S}'_i$ , являющийся простым циклом. Матрицу  $D'_i$  с помощью перестановок строк и столбцов можно привести к следующему верхне-блочно-диагональному виду

$$\begin{pmatrix} 1 & & * \\ & \ddots & \\ & & 1 \\ & & & D''_i \end{pmatrix},$$

где  $D''_i$  матрица, соответствующая графу  $\bar{S}''_i$ . Очевидно, что матрица  $D'_i$  имеет полный ранг тогда и только тогда, когда матрица  $D''_i$  имеет полный ранг.

Матрицу  $D''_i$  можно снова с помощью перестановки строк и столбцов преобразовать к матрице  $F$ , имеющей следующий вид

$$\begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & \ddots & \\ 1 & & & 1 & 1 \\ & & & & 1 \end{pmatrix},$$

где

$$f_{jl} = \begin{cases} 1, & \text{если } j = l, \\ 1, & \text{если } l = j + 1, \\ 1, & \text{если } j = m, l = 1, \\ 0, & \text{в противном случае.} \end{cases}$$

Пусть  $f_j = (f_{j1}, \dots, f_{jm})$ ,  $1 \leq j \leq m$ . Заменяя последнюю строку  $f_m$  матрицы  $F$  линейной комбинацией

$$f_m + \sum_{j=1}^{m-1} (-1)^j f_j,$$

мы получим в случае, если  $m$  нечетное число, невырожденную верхне-диагональную матрицу

$$\begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & \ddots & \\ & & & 1 & 1 \\ & & & & 2 \end{pmatrix}.$$

а в случае, если  $m$  четное число, вырожденную верхне-диагональную матрицу

$$\begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & \ddots & \\ & & & 1 & 1 \\ & & & & 0 \end{pmatrix}.$$

Таким образом, если каждый из графов  $\overline{S}_i$ ,  $i = 1, \dots, k$ , содержит нечетный цикл, то матрица  $D$  имеет полный ранг.

Пусть  $C = \langle Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}, Y_{t_1} \rangle$  – четный цикл. Определим линейную комбинацию  $w_1(C)$  строк матрицы  $W$ , соответствующих ребрам графа  $C$ :

$$w_1(C) = \sum_{i=1}^{k-1} w_{t_i, t_{i+1}} (-1)^{i+1} + w_{t_1, t_k} (-1)^{k+1}.$$

Пусть  $G$  подграф дополнительного графа, состоящий из двух простых нечетных циклов  $\langle Y_{t_1}, \dots, Y_{t_1} \rangle$  и  $\langle Y_{t_{i_2}}, \dots, Y_{t_{i_2}} \rangle$ , соединенных простой цепью  $\langle Y_{i_1}, \dots, Y_{i_2} \rangle$ . Определим линейную комбинацию  $w_2(G)$  строк матрицы  $W$  соответствующих ребрам графа  $G$ :

$$\begin{aligned} w_2(G) = & \sum_{j=1}^{i_1-1} w_{t_j, t_{j+1}} (-1)^{j+1} + w_{t_1, t_{i_1}} (-1)^{i_1+1} + \\ & 2(\sum_{j=i_1}^{i_2-1} w_{t_j, t_{j+1}} (-1)^{j+1}) + \\ & \sum_{j=i_2}^{n-1} w_{t_j, t_{j+1}} (-1)^{j+1} + w_{t_n, t_{i_2}} (-1)^{i_2}. \end{aligned}$$

Пусть дополнительный граф содержит четный цикл  $C_1$ . Преобразуем матрицу  $W$  и граф  $\bar{S}$  следующим образом. Удалим в графе  $\bar{S}$  произвольное ребро цикла  $C_1$ , а в матрице  $W$  – соответствующую этому ребру строку заменим линейной комбинацией  $w_1(C_1)$ . Будем повторять описанную операцию до тех пор, пока не получим граф  $\bar{S}^*(\bar{Y}, \bar{E}^*)$ , в котором нет четных циклов. Полученная в результате невырожденных линейных преобразований над строками матрицы  $W$  матрица  $W^*$  имеет следующий вид

$$W^* = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix},$$

где  $W_1 = (w_{ij})_{(Y_i, Y_j) \in \bar{E}^*}$  а  $W_2 = (0_n, F_1(C_i))_{i=1, \dots, r_1}$ .

Если граф  $\bar{S}^*$  не содержит циклов, то ранг матрицы  $W^*$  (а значит и матрицы  $W$ ) меньше  $n + 1$ , т.к. матрица  $W_1$  имеет не более  $n - 1$  строк.

Предположим, что граф  $\bar{S}^*$  содержит более чем  $n$  ребер. Тогда, т.к. граф  $\bar{S}^*$  связный, то он содержит подграф  $G_1$  состоящий из двух простых циклов, соединенных простой цепью. Преобразуем граф  $\bar{S}^*$  и матрицу  $W^*$  следующим образом. Удалим в графе  $\bar{S}^*$  любое ребро, принадлежащее одному из циклов графа  $G_1$ , а в матрице  $W^*$  соответствующую этому ребру строку заменим на линейную комбинацию  $w_2(G_1)$ . Будем повторять эту операцию пока не получим связный граф  $\bar{S}^{**}(\bar{Y}, \bar{E}^{**})$ , содержащий единственный нечетный цикл. В результате линейных операций над строками матрицы  $W^*$  мы получили матрицу  $W^{**}$ , имеющую следующий вид

$$W^{**} = \begin{pmatrix} W'_1 \\ W_2 \\ W_3 \end{pmatrix},$$

где  $W'_1 = (w_{ij})_{(Y_i, Y_j) \in \bar{E}^{**}}$ , а  $W_3 = (0_n, F_1(G_i))_{i=1, \dots, r_2}$ .

Пусть  $p$  – число компонент связности графа  $\bar{S}$ , которые не содержат нечетных циклов. Тогда, если условие 2 теоремы соблюдается, то ранг матрицы  $W$  равен  $n - p + 1$ , а если не соблюдается, то  $n - p$ .

Таким образом, ранг матрицы  $W^{**}$  (а значит и  $W$ ) равен  $n + 1$  тогда и только тогда, когда каждая компонента связности графа  $\bar{S}$  содержит по крайней мере один нечетный цикл и по крайне мере одна компонента содержит либо четный цикл  $C$  для которого

$F_1(C) \neq 0$ , либо подграф  $G$  состоящий из двух нечетных циклов, соединенных простой цепью для которого  $F_2(G) \neq 0$ .

Теорема доказана.

**Теорема 2.** *Пусть условия теоремы 1 соблюдаются. Если дополнительный граф является связным, то модель (1) является идентифицируемой.*

**ДОКАЗАТЕЛЬСТВО.** Как было отмечено при доказательстве теоремы 1, при выполнении условий теоремы параметры  $a_1, \dots, a_n, a$  определены с точностью до знака. Зафиксируем знаки  $a_1$  и  $a$ . Рассмотрим параметр  $a_i$ . Так как дополнительный граф является связным, то вершины  $Y_1$  и  $Y_i$  связаны простой цепью  $\langle Y_1, Y_{t_1}, \dots, Y_{t_i}, Y_i \rangle$ . Так как знаки  $\alpha_{1t_1}a_1a_{t_1}, \dots, \alpha_{t_i i}a_{t_i}a_i$  известны, то мы последовательно определяем знаки параметров  $a_{t_1}, \dots, a_{t_i}, a_i$ . Таким образом, при фиксированных знаках параметров  $a_1$  и  $a$ , знаки параметров  $a_i$ ,  $i = 2, \dots, n$ , определяются однозначно. Отсюда получаем, что знак произведения  $\alpha_{ij}a_ia_j$ ,  $i \neq j$ , определен однозначно. Используя (2), получаем, что параметры  $\{\omega_{ij}, Y_i - Y_j \notin \bar{E}\}$  определены однозначно. Теорема доказана.

#### 4. ПРИМЕРЫ ИДЕНТИФИЦИРУЕМЫХ МОДЕЛЕЙ

Пусть дополнительный граф структуры является связным и содержит граф  $G$ , представленный на рис. 1, тогда имеют место примеры 1 – 4.

**Пример 1.**

$$\{V_1, V_2, V_3\} \subset \mathbf{Y}^1, \quad \{V_4, V_5, V_6\} \subset \mathbf{Y}^2.$$

Легко видеть, что  $F_2(G) = 2$ .

**Пример 2.**

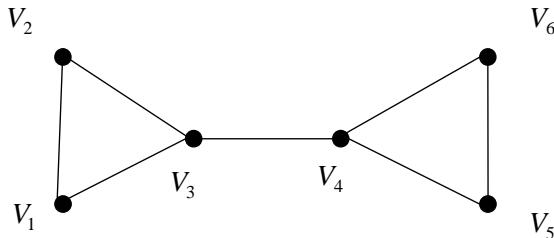
$$\{V_1, V_2, V_3, V_6\} \subset \mathbf{Y}^1, \quad \{V_4, V_5\} \subset \mathbf{Y}^2.$$

Легко видеть, что  $F_2(G) = 2$ .

**Пример 3.**

$$\{V_2, V_3\} \subset \mathbf{Y}^1, \quad \{V_1, V_4, V_5, V_6\} \subset \mathbf{Y}^2.$$

Легко видеть, что  $F_2(G) = 4$ .

Рис. 1. Граф  $G$ .**Пример 4.**

$$\{V_2, V_3, V_6\} \subset \mathbf{Y}^1, \quad \{V_1, V_4, V_5\} \subset \mathbf{Y}^2.$$

Легко видеть, что  $F_2(G) = 2$ .

Пусть дополнительный граф структуры является связным и содержит граф  $G'$ , представленный на рис. 2, тогда имеют место примеры 5,6.

**Пример 5.**

$$\{V_1, V_2\} \subset \mathbf{Y}^1, \quad \{V_3, V_4, V_5\} \subset \mathbf{Y}^2.$$

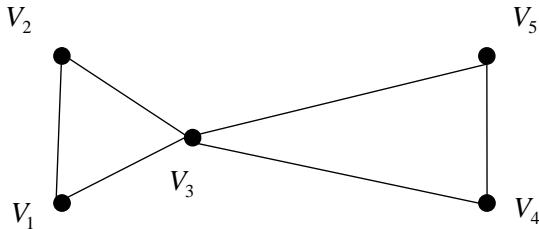
Легко видеть, что  $F_2(G') = 2$ .**Пример 6.**

$$\{V_1, V_2, V_5\} \subset \mathbf{Y}^1, \quad \{V_3, V_4\} \subset \mathbf{Y}^2.$$

Легко видеть, что  $F_2(G') = -2$ .

**Лемма 1.** Дополнительный граф леса с  $n$  вершинами, в котором нет вершины, имеющей  $n - 1$  смежную вершину, является связным.

**ДОКАЗАТЕЛЬСТВО.** При  $n = 2$  утверждение леммы очевидно. Предположим, что лемма справедлива при  $n \leq m$  и рассмотрим лес  $S = (\mathbf{Y}, \mathbf{E}_S)$  с  $m+1$  вершиной. Пусть  $Y' \in \mathbf{Y}$  произвольный лист леса  $S$  и пусть  $(Y', Y'') \in \mathbf{E}_S$ . Рассмотрим лес  $S' = (\mathbf{Y} \setminus Y', \mathbf{E}_S \setminus (Y', Y''))$ , полученный из  $S$  удалением ребра  $(Y', Y'')$ .

Рис. 2. Граф  $G'$ .

Предположим, что лес  $S'$  содержит вершину  $Y'''$  с  $m - 1$  смежной вершиной. В этом случае, легко видеть, что дополнительный граф  $\overline{S} = (\mathbf{Y}, \overline{\mathbf{E}}_S)$  является связным. Это следует из того, что

$$Y' - Y \in \overline{\mathbf{E}}_S, \quad Y \in \mathbf{Y} \setminus Y''$$

и

$$Y'' - Y \in \overline{\mathbf{E}}_S, \quad Y \in \mathbf{Y} \setminus (Y'', Y''').$$

Теперь предположим, что лес  $S'$  не содержит вершину с  $m - 1$  смежной вершиной. По предположению индукции, дополнительный граф  $\overline{S}' = (\mathbf{Y} \setminus Y', \overline{\mathbf{E}}_{S'})$  является связным. Но, так как

$$Y' - Y \in \overline{\mathbf{E}}_S, \quad Y \in \mathbf{Y} \setminus Y'',$$

то  $\overline{S}$  тоже является связным. Лемма доказана.

**Лемма 2.** *Среди любых пяти вершин дополнительного графа леса  $S = (\mathbf{Y}, \mathbf{E}_S)$  всегда найдутся три, соединенные между собой вершины.*

**ДОКАЗАТЕЛЬСТВО.** Рассмотрим подграф  $S' = (\{V_1, V_2, V_3, V_4, V_5\}, \mathbf{E}_{S'})$  графа  $S$ , причем для  $1 \leq i, j \leq 5$

$$(V_i, V_j) \in \mathbf{E}_{S'} \text{ тогда и только тогда, когда } (V_i, V_j) \in \mathbf{E}_S.$$

Утверждение леммы очевидно, если  $S'$  имеет не менее трех листьев.

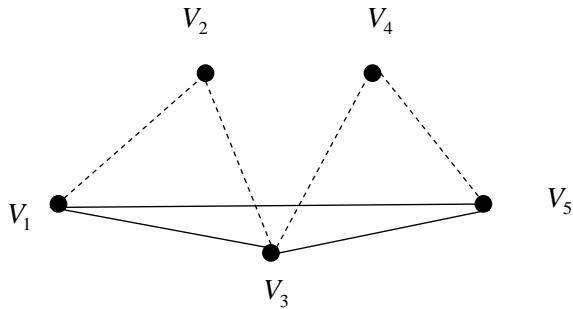


Рис. 3

Если \$S'\$ имеет только два листа, то тогда \$S'\$ является простой цепью и как нетрудно убедиться (см. рис. 3) и в этом случае утверждение леммы справедливо.

**Лемма 3.** Для произвольных четырех вершин \$V\_1, V\_2, V\_3, V\_4\$ дополнительного графа леса \$S = (\mathbf{Y}, \mathbf{E}\_S)\$

$$\{V_1 - V_2, V_2 - V_3, V_3 - V_4, V_4 - V_1\} \cap \overline{\mathbf{E}}_S \neq \emptyset.$$

**ДОКАЗАТЕЛЬСТВО.** Если утверждение леммы не выполняется, то

$$\{(V_1, V_2), (V_2, V_3), (V_3, V_4), (V_4, V_1)\} \subset \mathbf{E}_S,$$

а значит \$S\$ содержит цикл. Данное противоречие доказывает справедливость леммы.

**Теорема 3.** Пусть условия теоремы 1 выполняются и \$n\_1 \geq 4\$ и \$n\_2 \geq 4\$. Тогда, если график структуры \$S\$ является лесом, в котором нет вершины с \$n - 1\$ смежной вершиной, то модель будет идентифицируемой.

ДОКАЗАТЕЛЬСТВО. Рассмотрим индуцированный подграф дополнительного графа  $\bar{S}$  с множеством вершин  $\{V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8\}$ , причем  $\{V_1, V_2, V_3, V_4\} \subset \mathbf{Y}^1$  и  $\{V_5, V_6, V_7, V_8\} \subset \mathbf{Y}^2$ .

Рассмотрим множество из пяти вершин  $\{V_1, V_2, V_3, V_4, V_5\}$ . Согласно лемме 2 среди них есть три соединенных между собой вершины.

Предположим, что все три вершины принадлежат множеству  $\mathbf{Y}^1$ . Пусть для определенности это будут  $V_1, V_2$  и  $V_3$ . Среди множества вершин  $\{V_4, V_5, V_6, V_7, V_8\}$  снова найдутся три соединенные между собой вершины. Возможны два случая. В первом случае, все три вершины принадлежат множеству  $\mathbf{Y}^2$ , а во втором – две вершины принадлежат множеству  $\mathbf{Y}^2$ , а одна вершина множеству  $\mathbf{Y}^1$ . Используя лемму 3, получаем, что дополнительный граф структуры в первом случае, содержит подграф, рассмотренный в примере 1, а во втором случае – подграф, рассмотренный в примере 2.

Теперь предположим, что две вершины принадлежат множеству  $\mathbf{Y}^1$  (для определенности  $V_1$  и  $V_2$ ), а одна ( $V_5$ ) – множеству  $\mathbf{Y}^2$ . Тогда среди вершин множества  $\{V_5, V_6, V_7, V_8, V_3\}$  найдутся три соединенных между собой вершины. Снова возможны два случая. В первом случае, все три вершины принадлежат множеству  $\mathbf{Y}^2$ , а во втором – две вершины принадлежат множеству  $\mathbf{Y}^2$ , а одна вершина множеству  $\mathbf{Y}^1$ . Используя лемму 3, получаем, что дополнительный граф структуры в первом случае, содержит подграф, рассмотренный либо в примере 5, либо в примере 3, а во втором случае – подграф, рассмотренный либо в примере 4, либо в примере 6.

Так как, все графы, рассмотренные в примерах, соответствуют идентифицируемым моделям, теорема доказана.

## 5. УСЛОВИЯ ИДЕНТИФИЦИРУЕМОСТИ ПРИ ПРОИЗВОЛЬНОМ $k$

Для случая произвольного  $k$  нами получены достаточные условия идентифицируемости модели (1).

**Теорема 4.** Пусть  $a_i \neq 0$ ,  $i = 1, \dots, n$ ,  $a_{ij} \neq 0$ ,  $i, j = 1, \dots, k$ . Модель (1) локально идентифицируема, если

- (1) каждая компонента связности дополнительного графа  $\bar{S}$  содержит по крайней мере один нечетный цикл;

(2) для  $1 \leq i < j \leq k$  индуцированный подграф  $\overline{S}$  с множеством вершин  $\mathbf{Y}^i \cup \mathbf{Y}^j$  содержит либо четный цикл  $C$ , для которого  $F_1^{ij}(C) \neq 0$ , или подграф  $G$ , состоящий из двух простых нечетных циклов, соединенных простой цепью, для которого  $F_2^{ij}(G) \neq 0$ .

**ДОКАЗАТЕЛЬСТВО.** По аналогии с доказательством теоремы 1, мы будем использовать следующее представление для компонент матрицы ковариаций  $\Sigma = (\sigma_{ij})$  наблюдаемых случайных величин:

$$\sigma_{ij} = \begin{cases} \alpha_{ij}a_i a_j + \omega_{ij}, & \text{если } Y_i - Y_j \notin \overline{\mathbf{E}}, \\ \alpha_{ij}a_i a_j, & \text{если } Y_i - Y_j \in \overline{\mathbf{E}}, \end{cases} \quad (3)$$

где

$$\alpha_{ij} = \begin{cases} a_{lm}, & \text{если } Y_i \in \mathbf{Y}^l, Y_j \in \mathbf{Y}^m \text{ или} \\ & Y_j \in \mathbf{Y}^l, Y_i \in \mathbf{Y}^m, l \neq m, l, m = 1, \dots, k, \\ 1, & \text{в противном случае.} \end{cases}$$

Система равенств (2) задает отображение

$$\varphi : \Theta \rightarrow \mathbb{R}^{\frac{n(n+1)}{2}}.$$

Матрица Якоби  $J$ , соответствующая отображению, заданному системой равенств (2), имеет следующий блочно-верхне-диагональный вид

$$J = \begin{pmatrix} J_1 & * \\ & J(\alpha) \end{pmatrix},$$

где

$$J_1 = \left( \frac{\partial \sigma_{ij}}{\partial \omega_{ij}} \right)_{Y_i - Y_j \notin \overline{\mathbf{E}}} = I_{\frac{n(n+1)}{2} - |\overline{\mathbf{E}}|},$$

$$J(\alpha) = \left( \frac{\partial \sigma_{ij}}{\partial a_s}, s = 1, \dots, n, \frac{\partial \sigma_{ij}}{\partial a_{ml}}, 1 \leq m < l \leq k \right)_{Y_i - Y_j \in \overline{\mathbf{E}}}.$$

Легко видеть, что для  $i, j, Y_i - Y_j \in \overline{\mathbf{E}}, k = 1, \dots, n$ ,

$$\frac{\partial \sigma_{ij}}{\partial a_k} a_k = d_{ij}^k \alpha_{ij} a_i a_j, \quad \frac{\partial \sigma_{ij}}{\partial a} a_{ij} = h_{ij}^{ml} \alpha_{ij} a_i a_j,$$

где для  $k = 1, \dots, n$

$$d_{ij}^k = \begin{cases} 1, & \text{если } k = i, j, \\ 0, & \text{если } k \neq i, j, \end{cases}$$

и

$$h_{ij}^{ml} = \begin{cases} 1, & \text{если } \alpha_{ij} = a_{ml}, \\ 0, & \text{если } \alpha_{ij} = 1. \end{cases}$$

Пусть  $d_{ij} = (d_{ij}^1, \dots, d_{ij}^n)$  и  $h_{ij} = (h_{ij}^{ml}, 1 \leq m < l \leq k)$ ,  $w_{ij} = (d_{ij}, h_{ij})$ .

Образуем две матрицы  $D = (d_{ij})_{(i,j) \in \overline{E}}$  и  $W = (w_{ij})_{(i,j) \in \overline{E}}$ .

Очевидно, что

$$J(\alpha) \operatorname{diag}\{a_1, \dots, a_n, a\} = \operatorname{diag}\{\alpha_{ij} a_i a_j, Y_i - Y_j \in \overline{\mathbf{E}}\} W.$$

Ввиду того, что  $a_i \neq 0$ ,  $i = 1, \dots, n$ ,  $a_{ml} \neq 0$ ,  $1 \leq m < l \leq k$ , матрица  $J(\alpha)$  имеет полный ранг тогда и только тогда, когда матрица  $W$  имеет полный ранг.

Снова действуем по аналогии с доказательством теоремы 1. Пусть дополнительный граф содержит четный цикл  $C_1$ . Преобразуем матрицу  $W$  и граф  $\overline{S}$  следующим образом. Удалим в графе  $\overline{S}$  произвольное ребро цикла  $C_1$ , а в матрице  $W$  – соответствующую этому ребру строку заменим линейной комбинацией  $w_1(C_1)$ . Будем повторять описанную операцию до тех пор, пока не получим граф  $\overline{S}^*(\mathbf{Y}, \overline{\mathbf{E}}^*)$ , в котором нет четных циклов. Полученная в результате невырожденных линейных преобразований над строками матрицы  $W$  матрица  $W^*$  имеет следующий вид

$$W^* = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix},$$

где  $W_1 = (w_{ij})_{(Y_i, Y_j) \in \overline{\mathbf{E}}^*}$  а  $W_2 = (0_n, F_1(C_i))_{i=1, \dots, r_1}$  с  $F_1(C_i) = (F_1^{ml}(C_i), 1 \leq m < l \leq k)$ .

Предположим, что граф  $\overline{S}^*$  содержит более чем  $n$  ребер. Тогда он содержит подграф  $G_1$ , состоящий из двух простых циклов, соединенных простой цепью. Преобразуем граф  $\overline{S}^*$  и матрицу  $W^*$  следующим образом. Удалим в графе  $\overline{S}^*$  любое ребро, принадлежащее одному из циклов графа  $G_1$ , а в матрице  $W^*$  соответствующую этому ребру строку заменим на линейную комбинацию  $w_2(G_1)$ . Будем повторять эту операцию пока не получим связный граф  $\overline{S}^{**}(\mathbf{Y}, \overline{\mathbf{E}}^{**})$ , содержащий единственный нечетный цикл. В результате линейных операций над строками матрицы  $W^*$  мы

получили матрицу  $W^{**}$ , имеющую следующий вид

$$W^{**} = \begin{pmatrix} W'_1 \\ W_2 \\ W_3 \end{pmatrix},$$

где  $W'_1 = (w_{ij})_{(Y_i, Y_j) \in \overline{\mathcal{E}}^{**}}$ , а  $W_3 = (0_n, F_2(G_i))_{i=1, \dots, r_2}$  с  $F_2(G_i) = (F_2^{ml}(G_i), 1 \leq m < l \leq k)$ .

Пусть  $L_1 = (F_1(C_i))_{i=1, \dots, r_1}$ , а  $L_2 = (F_2(G_i))_{i=1, \dots, r_2}$ . Матрица  $W^{**}$  имеет блочно-верхне-диагональный вид

$$\begin{pmatrix} D^{**} & * \\ & L_1 \\ & L_2 \end{pmatrix}.$$

Таким образом, ранг матрицы  $W^{**}$  (а значит и  $W$ ) будет полным тогда и только тогда, когда ранг матрицы  $D^{**}$  равен  $n$ , а ранг матрицы  $\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$  равен  $\frac{k(k-1)}{2}$ .

Из доказательства теоремы 1 следует, что ранг матрицы  $D^{**}$  равен  $n$  тогда и только тогда, когда каждая компонента связности графа  $\overline{S}$  содержит по крайней мере один нечетный цикл.

Предположим, что условия 2 теоремы выполнены. Тогда, легкого видеть, что множество графов  $\{C_1, \dots, C_{r_1}, G_1, \dots, G_{r_2}\}$  содержит  $\frac{k(k-1)}{2}$  графов  $C_{ij}^{m_{ij}}$ ,  $1 \leq i < j \leq k$ ,  $m_{ij} = 1$  или  $m_{ij} = 2$ , таких, что  $F_{m_{ij}}^{ij}(C_{ij}^{m_{ij}}) \neq 0$ . Отсюда следует, что матрица  $\begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$  содержит подматрицу, которую перестановкой строк и столбцов можно привести к виду

$$\begin{pmatrix} F_{m_{12}}^{12}(C_{12}^{m_{12}}) & & & \\ & \ddots & & \\ & & F_{m_{k-1,k}}^{k-1,k}(C_{k-1,k}^{m_{k-1,k}}) & \end{pmatrix}.$$

Таким образом, при выполнении условий теоремы, ранг матрицы  $W$  будет полным. Теорема доказана.

**Теорема 5.** Пусть условия теоремы 4 соблюдаются. Если дополнительный граф структуры является связным, то модель (1) идентифицируема.

Доказательство теоремы аналогично доказательству теоремы 2.

**Теорема 6.** Пусть условия теоремы выполняются и  $n_i \geq 4$ ,  $i = 1, \dots, k$ . Тогда, если граф структуры  $S$  является лесом, в котором нет вершины с  $n - 1$  смежной вершиной, то модель будет идентифицируемой.

**ДОКАЗАТЕЛЬСТВО.** Из доказательства теоремы 3 следует, что в нашем случае условие 1 теоремы 4 соблюдается.

Для  $1 \leq i < j \leq k$  рассмотрим индуцированный подграф  $S_{ij} = (\mathbf{Y}^i \cup \mathbf{Y}^j, \mathbf{E}_{ij})$  графа  $S$ . Легко видеть, что условие 2 теоремы соблюдается, если граф  $S_{ij}$  содержит вершину с  $|\mathbf{Y}^i \cup \mathbf{Y}^j| - 1$  смежными вершинами. В тоже время из доказательства теоремы 3 следует, что условие 2 соблюдается и в случае, когда граф  $S_{ij}$  не содержит вершину с  $|\mathbf{Y}^i \cup \mathbf{Y}^j| - 1$  смежными вершинами. Теорема доказана.

## Литература

1. Стafeев C.B. К вопросу об оценивании параметров модели факторного анализа с зависимыми остатками // Труды ИПМИ КарНЦ РАН, 2003, с. 98-105.
2. Стafeев C.B. Факторный анализ с зависимыми остатками: проблема идентифицируемости и оценка параметров // Труды ИПМИ КарНЦ РАН, 2005, с. 119-130.
3. Spirtes P., Richardson T. Ancestral graph Markov models // Ann. Statist., 2002, v. 30, p. 962-1030.
4. Grzebyk M., Wild P., Chouaniere D. On identification of multi-factor model with correlated residuals // Biometrika, 2004, v. 91, 141-151.
5. Geiger D., Heckerman D., King H., Meek C. Stratified Exponential Families: Graphical Models and Model Selection // Ann. Statist., 2001, v. 29, p. 505-529.

УДК 512.56

## О ДЕЙСТВИИ СЛУЧАЙНОГО $s$ -ОТОБРАЖЕНИЯ КОНЕЧНОЙ БУЛЕВОЙ РЕШЕТКИ

В. М. ТАРАРИН, И. А. ЧЕПЛЮКОВА

$s$ -отображение конечной булевой решетки  $L$  – верхний эндоморфизм  $\varphi$  решетки  $L$  такой, что из  $h(x) \leq 1$  следует  $h(\varphi(x)) = h(x)$ , где  $h(x)$  – высота элемента  $x \in L$ . Изучается действие случайного  $s$ -отображения булевой решетки  $L_n$  длины  $n$  при  $n \rightarrow \infty$ .

Работа выполнена при поддержке программы "Алгебро-икомбинаторные и комбинаторные методы математической кибернетики" ОМН РАН и гранта 2005 года Фонда содействия отечественной науке.

Приведем ряд необходимых сведений. Понятия компоненты связности, цикла, циклических вершин графа отображения считаются известными [1], также как и понятие булевой решетки [2]. Операции пересечения и объединения в решетке  $L$  будем обозначать, соответственно, через  $\wedge$  и  $\vee$ , наименьший элемент (нуль) и наибольший элемент (единицу) в  $L$  будем обозначать, соответственно, через **0** и **1**. Высотой  $h(x)$  элемента  $x$  конечной булевой решетки  $L$  называется точная верхняя грань длин цепей  $\mathbf{0} = x_0 < x_1 < \dots < x_m = x$  между **0** и  $x$ . Имеем  $h(\mathbf{0}) = 0$ . Элементы  $x \in L$  такие, что  $h(x) = 1$  называются *атомами* решетки  $L$ . Длиной  $l(L)$  конечной булевой решетки  $L$  называется высота  $h(\mathbf{1})$  единичного элемента **1** решетки  $L$ . Известно [2], что конечная булева решетка однозначно, с точностью до изоморфизма, определяется своей длиной. Булеву решетку длины  $n$  будем обозначать через  $L_n$ . Известно [2], что  $|L_n| = 2^n$  и  $L_n$  имеет точно  $n$  атомов, множество которых будем обозначать через  $A_n$ . Идеалом решетки  $L_n$

называется непустое подмножество  $I \subseteq L_n$  такое, что для любых  $a, b \in I$  справедливо  $a \vee b \in I$ , и если  $x \leq a, a \in I$ , то  $x \in I$ . В решетке  $L_n$  любой идеал есть множество  $(a) = \{x \in L_n | \mathbf{0} \leq x \leq a\}$  для некоторого  $a \in L_n$ . Идеалы  $I$  решетки  $L_n$  также являются булевыми решетками, причем  $l(I) \leq n$ .

Отображение  $\varphi$  булевой решетки  $L_n$  в себя будем называть *s-отображением*, если  $\varphi$  – верхний эндоморфизм решетки  $L_n$ , т.е.  $\varphi(x \vee y) = \varphi(x) \vee \varphi(y)$  для любых  $x, y \in L$ , и для любого  $x \in L_n$  такого, что  $h(x) \leq 1$  справедливо равенство  $h(\varphi(x)) = h(x)$ .

Множество всех *s*-отображений решетки  $L_n$  будем обозначать через  $\Omega(L_n)$ . Очевидно, что  $\Omega(L_n)$  является полугруппой относительно суперпозиции отображений.

*Замечание.* Если  $\varphi$  – *s*-отображение булевой решетки  $L_n$ , то, в силу определения,  $\varphi(A_n) = \{\varphi(a) | a \in A_n\} \subseteq A_n$ , и, следовательно, ограничение  $\varphi$  на множество  $A_n$  атомов решетки  $L_n$  является отображением множества  $A_n$  в себя.

*Пример.* Пусть  $\mathcal{P}(X)$  – булева решетка всех подмножеств множества  $X$  относительно обычных теоретико-множественных операций объединения  $\bigcup$  и пересечения  $\bigcap$  и  $f$  – произвольное отображение множества  $X$  в себя. Определим отображение  $\varphi(f)$  множества  $\mathcal{P}(X)$  в себя, полагая  $\varphi(f)(Y) = f(Y)$ , где  $f(Y) = \{f(y) | y \in Y\}$ , для любого подмножества  $Y \subseteq X$ . Несложно показать, что  $\varphi(f)$  является *s*-отображением решетки  $\mathcal{P}(X)$ .

**Предложение 1.** Соответствие  $\pi$ , сопоставляющее *s*-отображению  $\varphi$  булевой решетки  $L_n$ , ограничение  $\overline{\varphi}$  отображения  $\varphi$  на множество  $A_n$  атомов решетки  $L_n$  является изоморфизмом полугруппы  $\Omega(L_n)$  на полугруппу  $\mathcal{T}(A_n)$  всех отображений множества  $A_n$  в себя.

**Доказательство.** В силу замечания, если  $\varphi \in \Omega(L_n)$ , то  $\pi(\varphi) = \overline{\varphi} \in \mathcal{T}(A_n)$ . Если  $\varphi_1, \varphi_2 \in \Omega(L_n)$  и  $\overline{\varphi}_1 = \overline{\varphi}_2$ , то  $\varphi_1(a) = \varphi_2(a)$  для любого  $a \in A_n$ . Любой элемент  $x \in L, x \neq \mathbf{0}$ , однозначно может быть представлен в виде объединения  $x = \bigvee a_i$  различных атомов  $a_1, \dots, a_k \in A_n$  [2], откуда получаем  $\varphi_1(x) = \varphi_1(a_1 \vee \dots \vee a_k) = \varphi_1(a_1) \vee \dots \vee \varphi_1(a_k) = \varphi_2(a_1) \vee \dots \vee \varphi_2(a_k) = \varphi_2(a_1 \vee \dots \vee a_k) = \varphi_2(x)$ , и, следовательно,  $\varphi_1 = \varphi_2$ . Поэтому  $\pi$  является инъективным отображением. Так же  $\pi$  – сюръективное отображение. Действитель-

но, если  $f \in \mathcal{T}(A_n)$ , то определим  $\varphi(f)$  по правилу:  $\varphi(f)(x) = \vee f(a_i)$ , где  $x = \vee a_i, a_1, \dots, a_k \in A_n$ . Несложно показать, что  $\varphi(f)$  является  $s$ -отображением решетки  $L_n$  и  $\overline{\varphi(f)} = f$ . Очевидно, что  $\pi(\varphi_1\varphi_2) = \pi(\varphi_1)\pi(\varphi_2)$  для любых  $\varphi_1, \varphi_2 \in \Omega(L_n)$ . Из выше изложенного вытекает справедливость предложения.

**Предложение 2.** *Если  $\varphi \in \Omega(L_n)$  и  $A_n$  – множество всех атомов булевой решетки  $L_n$ , то*

- (1) *равенство  $\varphi(x) = x, x \in L_n, x \neq 0$ , справедливо тогда и только тогда, когда множество  $A_n(x) = \{a \in A_n | a \leqslant x\}$  является объединением некоторых циклов графа отображения  $\varphi$  на множестве  $A_n$ ;*
- (2)  *$Fix_{L_n}(\varphi) = \{x \in L_n | \varphi(x) = x\}$  – булева решетка длины равной числу циклов (или, что эквивалентно, числу компонент связности) графа отображения  $\varphi$  на множестве  $A_n$ , причем, единичным элементом решетки  $Fix_{L_n}(\varphi)$  является элемент  $1^\varphi = \vee a_i$ , где  $a_i$  пробегает все циклические вершины графа отображения  $\varphi$  на множестве  $A_n$ ;*
- (3)  *$L_n^\varphi = (1^\varphi]$  – наибольший идеал решетки  $L_n$  такой, что ограничение отображения  $\varphi$  на множество  $L_n^\varphi$  является автоморфизмом решетки  $L_n^\varphi$ . Длина решетки  $L_n^\varphi$  равна числу циклических вершин графа отображения  $\varphi$  на множестве  $A_n$ ;*
- (4)  *$L_n^\varphi$  – множество всех циклических вершин графа отображения  $\varphi$  на множестве  $L_n$ .*

**ДОКАЗАТЕЛЬСТВО.** (1) Известно [2], что если  $x \in L_n$ , то  $x = \vee a_i$  для всех  $a_i \in A_n(x)$ , причем, элемент  $x$  однозначно определяется  $A_n(x)$ . Поэтому из  $\varphi(x) = x$  следует  $\varphi(A_n(x)) = \{\varphi(a) | a \in A_n(x)\} = A_n(x)$ , откуда вытекает, что  $A_n(x)$  является объединением некоторого числа циклов графа отображения  $\varphi$  на  $A_n$ . Обратное утверждение очевидно.

(2) Если  $x, y \in Fix_{L_n}(\varphi)$ , то, согласно пункту (1), множества  $A_n(x)$  и  $A_n(y)$  являются объединениями некоторых циклов графа отображения  $\varphi$  на  $A_n$ . Так как  $A_n(x \vee y) = A_n(x) \cup A_n(y)$  и  $A_n(x \wedge y) = A_n(x) \cap A_n(y)$  [2], то множества  $A_n(x \vee y)$  и  $A_n(x \wedge y)$  также являются объединением некоторых циклов графа отображения  $\varphi$  на  $L_n$ , откуда по (1) получаем, что  $x \vee y, x \wedge y \in Fix_{L_n}(\varphi)$ .

Таким образом,  $Fix_{L_n}(\varphi)$  – решетка,  $\mathbf{0} \in Fix_{L_n}(\varphi)$ . Очевидно, что единичным элементом решетки  $Fix_{L_n}(\varphi)$  является элемент  $\mathbf{1}^\varphi$  такой, что  $A_n(\mathbf{1}^\varphi)$  – объединение всех циклов графа отображения  $\varphi$  на  $A_n$ , т.е.  $A_n(\mathbf{1}^\varphi)$  – множество всех циклических вершин графа отображения  $\varphi$  на множестве  $A_n$ . Пусть  $x \in Fix_{L_n}(\varphi), x \neq \mathbf{0}, x \neq \mathbf{1}^\varphi$ . Существует элемент  $y \in L_n$  такой, что  $A_n(y)$  – объединение всех циклов графа отображения  $\varphi$  на  $A_n$ , не входящих во множество  $A_n(x)$ . Тогда, очевидно,  $y \in Fix_{L_n}(\varphi)$  и  $x \wedge y = \mathbf{0}, x \vee y = \mathbf{1}^\varphi$ . Таким образом,  $Fix_{L_n}(\varphi)$  – булева решетка. Очевидно, что атомами решетки  $Fix_{L_n}(\varphi)$  являются элементы  $x$  такие, что  $A_n(x)$  – цикл графа отображения  $\varphi$  на  $A_n$ , откуда следует, что  $l(Fix_{L_n}(\varphi))$  – число циклов графа отображения  $\varphi$  на  $L_n$ .

(3) Пусть  $x \in L_n$  и  $(x)$  – идеал такой, что ограничение отображения  $\varphi$  на  $(x)$  является автоморфизмом решетки  $(x)$ . Тогда, в частности,  $\varphi((x)) = (x)$ . Покажем, что  $\varphi(x) = x$ . Предположим, что  $\varphi(x) \neq x$ . Так как  $\varphi(x) \in (x)$ , то  $\varphi(x) < x$ . Так как  $\varphi((x)) = (x)$ , то существует  $y \in (x)$  такой, что  $\varphi(y) = x$ . Так как  $y \leqslant x$ , то  $y \vee x = x$ , откуда  $\varphi(x) = \varphi(y \vee x) = \varphi(y) \vee \varphi(x) = x \vee \varphi(x) = x$ , что противоречит неравенству  $\varphi(x) < x$ . Поэтому  $\varphi(x) = x$  и  $x \in Fix_{L_n}(\varphi)$ . По пункту (1)  $A_n(x)$  – объединение некоторых циклов графа отображения  $\varphi$  на  $A_n$  и, следовательно, ограничение  $\varphi$  на  $A_n(x)$  является подстановкой множества  $A_n(x)$ . Так как  $A_n(x)$  – атомы решетки  $(x)$  и  $\varphi$  – подстановка множества  $A_n(x)$ , то, как несложно показать, ограничение  $\varphi$  на множество  $(x)$  является автоморфизмом решетки  $(x)$ . Из выше изложенного вытекает, что  $L_n^\varphi = (\mathbf{1}^\varphi]$  – наибольший идеал решетки  $L_n$  такой, что ограничение  $\varphi$  на  $L_n^\varphi$  является автоморфизмом решетки  $L_n^\varphi$ , и, по пункту(2), длина решетки  $L_n^\varphi$  равна числу циклических вершин графа отображения  $\varphi$  на множество  $A_n$ .

(4) В силу пункта (3), если  $x \in L_n^\varphi$ , то  $x$  – циклическая вершина графа отображения  $\varphi$  на множестве  $A_n$ . Пусть  $x$  – циклическая вершина графа  $\varphi$  на  $L_n$ . Тогда существует наименьшее натуральное число  $k$  такое, что  $\varphi^k(x) = x$ . Пусть  $y = \vee \varphi^i(x), i \in \{1, \dots, k\}$ . Тогда  $\varphi(y) = \varphi(\vee \varphi^i(x)) = \vee \varphi^{i+1}(x) = y$ , и, следовательно,  $y \in Fix_{L_n}(\varphi)$ , откуда в силу  $x \leqslant y \leqslant \mathbf{1}^\varphi$  получаем, что  $x \in L_n^\varphi$ .

Предложение доказано.

Введем на множестве  $\Omega(L_n)$  всех  $s$ -отображений булевой решетки  $L_n$  равномерное распределение вероятностей.

**Теорема 1.** Пусть  $\nu_n(\varphi)$  – длина решетки  $Fix_{L_n}(\varphi)$  случайного  $s$ -отображения  $\varphi \in \Omega(L_n)$ . Тогда при  $n \rightarrow \infty$

$$\mathbf{P}\{\nu_n(\varphi) = N\} = \frac{1}{\sqrt{\pi \ln n}} \exp \left\{ -\frac{(N - \frac{1}{2} \ln n)^2}{\ln n} \right\} (1 + o(1))$$

равномерно относительно целых  $N$ , для которых  $(N - \frac{1}{2} \ln n)(\ln n)^{-7/12}$  лежит в любом конечном интервале, и

$$\mathbf{E} \nu_n(\varphi) = \frac{1}{2} \ln n + o(\ln n).$$

**Теорема 2.** Пусть  $\nu_n(\varphi)$  – длина решетки  $Fix_{L_n}(\varphi)$  случайного  $s$ -отображения  $\varphi \in \Omega(L_n)$ . Тогда при  $n \rightarrow \infty$

$$\frac{(m-1)! \sqrt{n}}{(\ln \sqrt{n})^{m-1}} \mathbf{P}\{\nu_n(\varphi) = m\} = W + o(1),$$

где

$$W = \begin{cases} \sqrt{\pi/2}, & m = o(\ln n), \\ (2^\gamma \gamma \Gamma(\gamma))^{-1}, & m = \gamma \ln n + o(\ln n), 0 < \gamma < \infty. \end{cases}$$

**Теорема 3.** Пусть  $\nu_n(\varphi)$  – длина решетки  $Fix_{L_n}(\varphi)$  случайного  $s$ -отображения  $\varphi \in \Omega(L_n)$  и  $n \rightarrow \infty, \beta \rightarrow \infty, m/\ln n = O(\ln n)$ . Тогда

$$\frac{(m-1)! \sqrt{2\beta n}}{(-\ln(1-x_1))^{m-1} e^{\beta/2} \beta^\mu} \mathbf{P}\{\nu_n(\varphi) = m\} = \exp\{-\gamma_1\} (1 + o(1)),$$

где

$$x_1 = 1 - \frac{1}{\alpha \ln \alpha}, \gamma_1 = \frac{\beta}{8 \ln \alpha} \ln^2(2\mu), \alpha = \frac{\sqrt{n}}{m},$$

$$\beta = \frac{m}{\ln \alpha}, \mu = \frac{\beta}{2} \left( 1 - \frac{\ln \ln \alpha}{\ln \alpha} \right).$$

**Теорема 4.** Пусть  $\lambda^{(n)}(\varphi)$  – длина идеала  $L_n^\varphi$  случайного  $s$ -отображения  $\varphi \in \Omega(L_n)$ . Тогда для любого  $N = 1, 2, \dots, n$

$$\mathbf{P}\{\lambda^{(n)}(\varphi) = N\} = \frac{N(n-1)!}{n^N (n-N)!}.$$

Обозначим через  $\Omega_m(L_n)$  множество всех  $s$ -отображений  $\varphi \in \Omega_m(L_n)$  таких, что  $l(Fix_{L_n}(\varphi)) = m$ , и введем на  $\Omega_m(L_n)$  равномерное распределение вероятностей.

**Теорема 5.** Пусть  $\lambda^{(n,m)}(\varphi)$  – длина идеала  $L_n^\varphi$  случайного  $s$ -отображения  $\varphi \in \Omega_m(L_n)$ . Если  $n \rightarrow \infty, m/\ln n \rightarrow 0$ , то равномерно относительно целых положительных чисел  $N$ , для которых  $z = N/\ln n$  лежит в любом конечном интервале

$$\sqrt{n} \mathbf{P} \left\{ \lambda^{(n,m)}(\varphi)/\sqrt{n} = z \right\} = \sqrt{2/\pi} e^{-z^2/2} + o(1),$$

$$\mathbf{E} \lambda^{(n,m)}(\varphi) = \sqrt{\frac{2n}{\pi}} + o(\sqrt{n}).$$

**Теорема 6.** Пусть  $\lambda^{(n,m)}(\varphi)$  – длина идеала  $L_n^\varphi$  случайного  $s$ -отображения  $\varphi \in \Omega_m(L_n)$ . Если  $n, m \rightarrow \infty, m/\ln n \rightarrow \gamma$ , где  $\gamma$  – некоторая положительная постоянная, то равномерно относительно целых положительных  $N$ , для которых  $z = N/\ln n$  лежит в любом конечном интервале

$$\sqrt{n} \mathbf{P} \left\{ \lambda^{(n,m)}(\varphi) = z \right\} = \frac{2^\gamma}{\sqrt{2\pi}} \frac{\Gamma(\gamma)}{\Gamma(2\gamma)} z^{2\gamma} e^{-z^2/2} + o(1),$$

$$\mathbf{E} \lambda^{(n,m)}(\varphi) = \frac{\sqrt{2\pi}\Gamma(\gamma+1)}{\Gamma(\gamma+1/2)} (1 + o(1)).$$

**Теорема 7.** Пусть  $\lambda^{(n,m)}(\varphi)$  – длина идеала  $L_n^\varphi$  случайного  $s$ -отображения  $\varphi \in \Omega_m(L_n)$ . Если  $n \rightarrow \infty, m/\ln n \rightarrow \infty, m/\ln n = O(\ln n), z = N/\sqrt{n}$ , то равномерно относительно целых положительных  $N$ , для которых  $z - \sqrt{2\mu}$  лежит в любом фиксированном конечном интервале

$$\beta^\mu e^{-\beta/2} \sqrt{\pi n} \mathbf{P} \left\{ \lambda^{(n,m)}(\varphi) = N \right\} = z^{2\mu} e^{-z^2/2} (1 + o(1)),$$

где

$$\alpha = \frac{\sqrt{n}}{m}, \beta = \frac{m}{\ln \alpha}, \mu = \frac{\beta}{2} \left( 1 - \frac{\ln \ln \alpha}{\ln \alpha} \right).$$

**Теорема 8.** Пусть  $\lambda^{(n,m)}(\varphi)$  – длина идеала  $L_n^\varphi$  случайного  $s$ -отображения  $\varphi \in \Omega_m(L_n)$ . Если  $n, m \rightarrow \infty$  так, что  $m = n - l, l = o(\sqrt{n})$ ,

то для целых положительных  $N = n - j$ ,  $0 \leq j \leq l$ , справедливо

$$\mathbf{P} \left\{ \lambda^{(n,m)}(\varphi) = N \right\} = C_l^j \frac{2^j}{3^l} (1 + o(1)),$$

$$\mathbf{E} \lambda^{(n,m)}(\varphi) = \frac{2l}{3} + o(l).$$

**ДОКАЗАТЕЛЬСТВО.** Пусть на множестве  $\Omega(L_n)$  задано равномерное распределение вероятностей, тогда в силу предложения 1 на полугруппе  $\mathcal{T}(A_n)$  задано равномерное распределение вероятностей. Справедливость теорем 1 – 8 вытекает из пункта (2) предложения 2 и результатов, соответственно, работ [1, 3, 4]. Теорема 4 следует из пункта (3) предложения 2 и работы [5]. Теоремы 5 – 8 вытекают из пунктов (2) и (3) предложения 2 и результатов, соответственно работ [3, 4, 6].

## Литература

1. Колчин В. Ф. Случайные отображения. М.: Наука, 1984.
2. Биркгоф Г. Теория решеток. М.: Наука, 1984.
3. Павлов Ю. Л. Случайные леса. Петрозаводск, 1996.
4. Чеплюкова И. А. Один случай предельного распределения числа циклических вершин в случайному отображении // Дискретная математика. 2004, т. 16, вып. 3, с. 76–84.
5. Harris B. Probability distributions related to random mappings. Ann. Math. Statist. 1960. Vol.31, issue 4, p.1045-1062.
6. Чеплюкова И. А. О распределении числа циклических вершин в случайному отображении с известным числом циклов  
Труды ИПМИ КарНЦ РАН, Петрозаводск, 2005, вып.6, с.142-151.

УДК 519.833

**ИГРА НАИЛУЧШЕГО ВЫБОРА С ВОЗМОЖНОСТЬЮ ОТКАЗА  
ОТ ПРЕДЛОЖЕНИЯ И ПЕРЕРАСПРЕДЕЛЕНИЕМ  
ВЕРОЯТНОСТЕЙ**

А. А. ФАЛЬКО

Рассматривается игра  $m$  лиц наилучшего выбора с полной информацией. Найдены оптимальные стратегии для случая возможности отказа от предложения и перераспределения вероятностей.

**1. ВВЕДЕНИЕ**

В данной работе рассматривается игра  $m$  лиц наилучшего выбора. Пусть имеется  $m$  предприятий, каждое из которых хочет принять на работу по специалисту в одной и той же области. Всего имеется  $n$  претендентов на свободное место и квалификация каждого равномерно распределена на отрезке  $[0,1]$ . Начальники предприятий (игроки) по очереди беседуют с каждым претендентом и после этого выносят решение принять его или отвергнуть. Если  $i$ -ый игрок решает предложить претенденту работу, то претендент может согласиться принять предложение с вероятностью  $p_i$ ,  $i = \overline{1, m}$ ,  $p_1 + p_2 + \dots + p_m \leq 1$ . Когда некоторые начальники уже выбрали специалистов, вероятности, с которыми остальные претенденты могли бы согласиться работать на этих предприятиях, перераспределяются на оставшиеся. Если все игроки отвергают текущего претендента, то рассматривается следующий, причем, отвергнув претендента, к нему нельзя будет вернуться в дальнейшем. Если никто из претендентов не был принят, то предприятие терпит убытки  $C$  из-за нехватки специалистов,  $C \in [0, 1]$ .

Существуют различные варианты критериев в задаче наилучшего выбора: 1) максимизация вероятности найти наилучшего претендента, 2) максимизация математического ожидания выигрыша, 3) минимизация ожидаемого абсолютного ранга претендента. Сакагути и Мазалов [1] рассмотрели сценарий задачи для двух игроков, совместно выбирающих претендента, в котором минимизируется ожидаемый абсолютный ранг претендента, и нет никакой информации о поступающих претендентах. Тамаки [2] исследовал вариант задачи с возможностью отказа претендента от предложения и получил формулу для нахождения выигрыша, когда число претендентов стремится к бесконечности. Задача максимизации математического ожидания выигрыша часто называется задачей о продаже недвижимости. Фергюсон [3] рассмотрел вариант данной задачи для нескольких игроков. Случай максимизации вероятности найти наилучшего претендента для двух и трех игроков, когда приоритет игрока определяется с заданной вероятностью, изучен Сакагути [4]. Гарнаев и Соловьев [5] рассмотрели сценарий для двух игроков и возможности отказа от предложения, а также сценарий игры с пороговым значением для квалификации претендента.

В данной статье рассматривается случай  $m$  игроков с возможностью отказа от предложения и перераспределением вероятностей. Необходимо максимизировать вероятность найти наилучшего претендента.

## 2. Случай с одним и двумя игроками

Сначала рассмотрим случай с одним игроком. Каждый претендент имеет квалификацию  $x$ , которая равномерно распределена на отрезке  $[0, 1]$ . Вероятность того, что претендент согласится принять предложение равна  $p$ ,  $p \leq 1$ . Тогда  $\bar{p} = 1 - p$  – вероятность отказа от предложения. Решение такой задачи получено Гарнаевым и Соловьевым [5] для случая двух игроков и возможности отказа от предложения. Если на шаге  $r$  игрок отказывается от претендента, то рассматривается игра, в которой требуется рассмотреть  $r-1$  оставшихся претендентов. Обозначим  $u_{1,r}(p)$  – ожидаемый выигрыш игрока в этом случае.

Если  $r = 1$ , то игрок должен предложить претенденту работу

независимо от его квалификации. Тогда, учитывая, что  $x$  равномерно распределена на  $[0, 1]$ , получим

$$u_{1,1}(p) = \int_0^1 px \, dx + \int_0^1 \bar{p}(-C)dx = \frac{p}{2} - \bar{p}C.$$

Если же имеются  $r - 1$  претендентов, то

$$u_{1,r}(p) = \max_{\varphi} \left\{ \int_0^1 \left[ px\varphi(x) + \bar{p}u_{1,r-1}(p)\varphi(x) + u_{1,r-1}(p)\overline{\varphi(x)} \right] dx \right\},$$

где  $\varphi(x)$  – вероятность принять претендента с квалификацией  $x$ ,  $\overline{\varphi(x)} = 1 - \varphi(x)$ .

Таким образом, если имеются  $r - 1$  претендентов, то игрок примет текущего претендента с квалификацией  $x$ , тогда и только тогда, когда  $x \geq u_{1,r-1}(p)$ .

Следовательно,

$$\begin{aligned} u_{1,r}(p) &= \int_0^{u_{1,r-1}(p)} u_{1,r-1}(p) \, dx + \int_{u_{1,r-1}(p)}^1 (px + \bar{p}u_{1,r-1}(p)) \, dx \\ &= \frac{p}{2}(1 - u_{1,r-1}(p))^2 + u_{1,r-1}(p). \end{aligned}$$

Получим выражение

$$u_{1,r}(p) = \frac{p}{2}(1 - u_{1,r-1}(p))^2 + u_{1,r-1}(p). \quad (1)$$

Следует отметить, что  $\lim_{r \rightarrow \infty} u_{1,r}(p) = 1$  и этот предел не зависит от  $p$  и  $C$ .

Если на каком-нибудь шаге  $u_{1,r-1}(p) < 0$ , то

$$u_{1,r}(p) = \frac{p}{2} + \bar{p}u_{1,r-1}(p).$$

Рассмотрим случай двух предприятий. Назовем начальника первого предприятия Игроком 1, а второго – Игроком 2. Вероятность того, что претендент согласится принять предложение работать на предприятии  $i$ , равна  $p_i$ , где  $i = 1, 2$ ,  $p_1 + p_2 = p \leq 1$ ,  $\bar{p} = 1 - p$ .

Первоначально рассмотрим ситуацию, когда один игрок уже нанял на работу специалиста, и другому осталось рассмотреть  $r$  претендентов. В этом случае его выигрыш равен  $u_{1,r}(p)$ .

Теперь можно перейти к случаю двух игроков и  $r$  претендентов. Обозначим ожидаемый выигрыш  $i$ -го игрока  $u_{2,r}^i$ ,  $i = 1, 2$ . Если  $r = 1$ , то оба игрока заинтересованы принять претендента, так как они не хотят терпеть убытки из-за нехватки специалистов на своем предприятии. Следовательно, ожидаемые выигрыши каждого из игроков равны  $u_{2,1}^i = u_{1,1}(p_i)$ . Так как  $u_{2,1}^i(p_i)$  – возрастающие функции, то чем выше вероятность того, что претендент согласится работать на предприятии  $i$ , тем больше выигрыш  $i$ -го игрока.

Рассмотрим случай, когда имеются  $r$  претендентов. На первом шаге игру можно описать биматрицей  $M_r^2(x)$ .

$$M_r^2(x) = \begin{array}{cc} \Pi & \text{O} \\ \text{O} & \end{array} \left( \begin{array}{cc} (m_{11}^1, m_{11}^2) & (m_{12}^1, m_{12}^2) \\ (m_{21}^1, m_{21}^2) & (m_{22}^1, m_{22}^2) \end{array} \right),$$

где

$$\begin{aligned} m_{11}^1 &= p_1 x + p_2 u_{1,r-1} + \bar{p} u_{2,r-1}^1; \\ m_{11}^2 &= p_2 x + p_1 u_{1,r-1} + \bar{p} u_{2,r-1}^2; \\ m_{12}^1 &= p_1 x + \bar{p}_1 u_{2,r-1}^1; \\ m_{12}^2 &= p_1 u_{1,r-1} + \bar{p}_1 u_{2,r-1}^2; \\ m_{21}^1 &= p_2 u_{1,r-1} + \bar{p}_2 u_{2,r-1}^1; \\ m_{21}^2 &= p_2 x + \bar{p}_2 u_{2,r-1}^2; \\ m_{22}^1 &= u_{2,r-1}^1; \\ m_{22}^2 &= u_{2,r-1}^2. \end{aligned}$$

Стратегия Принять доминирует стратегию Отклонить для Игрока 1 тогда и только тогда, когда  $x \geq u_{2,r-1}^1$ , и стратегия Принять доминирует стратегию Отклонить для Игрока 2 тогда и только тогда, когда  $x \geq u_{2,r-1}^2$ . Не ограничивая общности, можно считать, что  $p_1 \geq p_2$ , тогда  $u_{2,r-1}^1 \geq u_{2,r-1}^2$ . Отсюда, если  $x \geq u_{2,r-1}^1$ , то для обоих игроков стратегия Принять доминирует стратегию Отклонить. При  $x \in [u_{2,r-1}^2, u_{2,r-1}^1]$  стратегия Принять доминирует стратегию Отклонить для Игрока 2 и Стратегия Отклонить доминирует стратегию Принять для Игрока 1. При  $x < u_{2,r-1}^2$  стратегия Отклонить доминирует стратегию Принять для обоих

игроков. Таким образом, оптимальной стратегией Игрока 1 является принятие претендента тогда и только тогда, когда  $x \geq u_{2,r-1}^1$ , и оптимальной стратегией Игрока 2 является принятие претендента тогда и только тогда, когда  $x \geq u_{2,r-1}^2$ . Учитывая, что  $x$  равномерно распределена на  $[0,1]$ , находим

$$\begin{aligned} u_{2,r}^1 &= \int_0^{u_{2,r-1}^2} u_{2,r-1}^1 dx + \int_{u_{2,r-1}^2}^{u_{2,r-1}^1} (p_2 u_{1,r-1} + \bar{p}_2 u_{2,r-1}^1) dx \\ &\quad + \int_{u_{2,r-1}^1}^1 (p_1 x + p_2 u_{1,r-1} + \bar{p} u_{2,r-1}^1) dx; \\ u_{2,r}^2 &= \int_0^{u_{2,r-1}^2} u_{2,r-1}^2 dx + \int_{u_{2,r-1}^2}^{u_{2,r-1}^1} (p_2 x + \bar{p}_2 u_{2,r-1}^2) dx \\ &\quad + \int_{u_{2,r-1}^1}^1 (p_2 x + p_1 u_{1,r-1} + \bar{p} u_{2,r-1}^2) dx. \end{aligned}$$

Вычисляя интегралы, получим выражения для  $u_{2,r}^1$  и  $u_{2,r}^2$ ,

$$u_{2,r}^1 = \frac{p_1}{2}(1 - u_{2,r-1}^1)^2 + u_{2,r-1}^1 + p_2(1 - u_{2,r-1}^2)(u_{1,r-1} - u_{2,r-1}^1);$$

$$u_{2,r}^2 = \frac{p_2}{2}(1 - u_{2,r-1}^2)^2 + u_{2,r-1}^2 + p_1(1 - u_{2,r-1}^1)(u_{1,r-1} - u_{2,r-1}^2).$$

Учитывая, что  $\lim_{r \rightarrow \infty} u_{1,r} = 1$ , получим  $\lim_{r \rightarrow \infty} u_{2,r}^i = 1$ ,  $i = 1, 2$ , который также не зависит от  $C$  и  $p_i$ .

### 3. Случай $m$ ИГРОКОВ

Рассуждая аналогично, данный сценарий можно рассмотреть и для  $m$  предприятий. Пусть имеется  $m$  предприятий и  $r$  претендентов. Вероятность того, что претендент согласится принять предложение работать на предприятии  $i$ , равна  $p_i$ , где  $i = \overline{1, m}$  и  $\sum_{i=1}^m p_i = p \leq 1$ . Обозначим ожидаемый выигрыш  $i$ -го игрока  $u_{m,r}^i$ ,  $i = \overline{1, m}$ . Если  $r = 1$ , то все игроки заинтересованы принять претендента, так как они не хотят терпеть убытки. Следовательно, здесь имеется единственное равновесие по Нэшу и ожидаемые выигрыши каждого из игроков равны  $u_{m,1}^i = u_{1,1}(p_i)$ . Так как  $u_{m,1}^i$  — возрастающие функции, то чем выше вероятность того, что

претендент согласится работать на предприятии  $i$ , тем больше выигрыш  $i$ -го игрока.

Перейдем к случаю  $r$  претендентов. Не ограничивая общности, можно считать, что

$p_1 \geq p_2 \geq \dots \geq p_m$ , тогда  $u_{m,r-1}^1 \geq u_{m,r-1}^2 \geq \dots \geq u_{m,r-1}^m$ .

Выигрыш  $i$ -го игрока вычисляется по формуле

$$u_{m,r}^i = \frac{p_i}{2}(1 - u_{m,r-1}^i)^2 + u_{m,r-1}^i + \sum_{j=1, j \neq i}^m [p_j(1 - u_{m,r-1}^j)(u_{m-1,r-1}^i - u_{m,r-1}^i)],$$

$$i = \overline{1, m},$$

(2)

где

$$u_{m,r}^i = u_{m,r}^i(p_1, \dots, p_m),$$

$$u_{m,r-1}^i = u_{m,r-1}^i(p_1, \dots, p_m), \quad u_{m,r-1}^j = u_{m,r-1}^j(p_1, \dots, p_m),$$

$$u_{m-1,r-1}^i = u_{m-1,r-1}^i\left(p_1 + \frac{p_j}{m-1}, \dots, p_{j-1} + \frac{p_j}{m-1}, p_{j+1} + \frac{p_j}{m-1}, \dots, p_m + \frac{p_j}{m-1}\right).$$

Таким образом, решение задачи для  $m$  игроков находится по формулам (1), (2).

Рассмотрим частный случай данной задачи. Пусть для претендентов все предприятие равнозначны, то есть  $p_1 = p_2 = \dots = p_m = \frac{p}{m}$ . Тогда выигрыш каждого игрока вычисляется по формуле

$$\begin{aligned} u_{m,r} &= \int_0^{u_{m,r-1}} u_{m,r-1} dx + \int_{u_{m,r-1}}^1 \left[ \frac{p}{m}x + \frac{(m-1)p}{m}u_{m-1,r-1} + (1-p)u_{m,r-1} \right] dx \\ &= \frac{p}{2m}(1 - u_{m,r-1})^2 + u_{m,r-1} + \frac{(m-1)p}{m}(1 - u_{m,r-1})(u_{m-1,r-1} - u_{m,r-1}), \end{aligned} \quad (3)$$

где

$$u_{m,r} = u_{m,r}(\frac{p}{m}),$$

$$u_{m,r-1} = u_{m,r-1}(\frac{p}{m}),$$

$$u_{m-1,r-1} = u_{m-1,r-1}(\frac{p}{m-1}).$$

Тогда  $\lim_{r \rightarrow \infty} u_{m,r} = 1$  и  $\lim_{m \rightarrow \infty} u_{m,r} = 0$ .  
Если на каком-нибудь шаге  $u_{m,r-1} < 0$ , то

$$\begin{aligned} u_{m,r} &= \int_0^1 \left[ \frac{p}{m}x + \frac{(m-1)p}{m}u_{m-1,r-1} + (1-p)u_{m,r-1} \right] dx \\ &= \frac{p}{2m} + \frac{(m-1)p}{m}u_{m-1,r-1} + (1-p)u_{m,r-1}. \end{aligned} \quad (4)$$

При  $0 \leq C \leq \frac{p}{2(m-p)}$  все функции выигрышей неотрицательны, и следовательно вычисляются только по формуле (3).

В таблице 1 приводятся численные результаты для различных  $m$  и  $p$  при  $C = 0$  и  $C = 0.5$ . На рис. 1 показаны функции выигрыша  $u_{5,10}$  для различных значений  $C$ .

ТАБЛИЦА 1. Значение выигрыша  $u_{m,10}$  для  $C = 0$  и  $C = 0.5$

$m \setminus p$	0.25	0.5	1	0.25	0.5	1
1	0.577	0.741	0.861	0.510	0.722	0.861
2	0.461	0.659	0.814	0.359	0.630	0.815
3	0.368	0.580	0.767	0.220	0.537	0.768
4	0.296	0.508	0.720	0.098	0.446	0.719
5	0.244	0.443	0.672	-0.005	0.357	0.670
6	0.205	0.387	0.625	-0.084	0.272	0.619
7	0.177	0.340	0.579	-0.143	0.192	0.565
8	0.155	0.301	0.534	-0.187	0.120	0.510
9	0.138	0.270	0.492	-0.222	0.056	0.454
10	0.124	0.245	0.454	-0.250	0	0.398

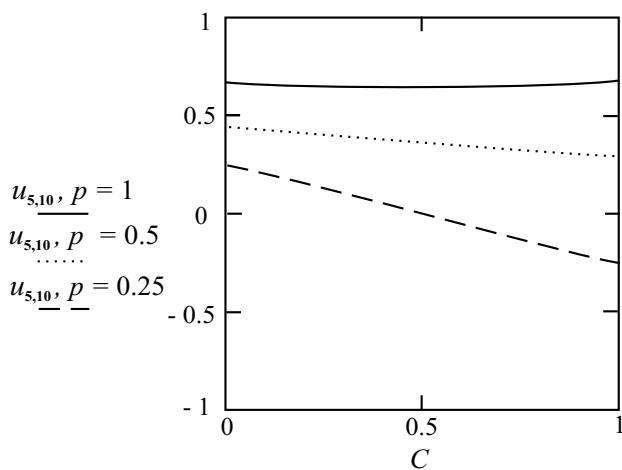


Рис. 1

На Рис.1 видно, что при  $p = 1$  выигрыши игроков  $u_{5,10}$  почти не зависят от значения  $C$ .

### Литература

1. *Sakaguchi M., Mazalov V.* A non-zero-sum no-information best-choice game // Mathematical Methods of Operation Research, 2004, Vol.60, pp. 437-451.
2. *Tamaki M.* Minimal expected ranks for the secretary problems with uncertain selection // Game Theory, Optimal Stopping, Probability and Statistics, ed. Bruss F. T. and Cam L. Le, Institute of Mathematical Statistics, 2000, pp.127-139.
3. *Ferguson T.* Selection by committee // Annals of the International Society of Dynamic Games, Vol.7, Advances in Dynamic Games Application to Economics, Finance, Optimization and Stochastic Control, 2005, pp. 203-209.
4. *Sakaguchi M.* Optimal stopping games where players have weighted privilege // Annals of the International Society of Dynamic Games, Vol.7, Advances in Dynamic Games Application to Economics, Finance, Optimization and Stochastic Control, 2005, pp. 285-294.
5. *Garnaev A., Solovyev A.* On a two department multi stage game // International Workshop Optimal Stopping and Stochastic Control, august 22-26, 2005, Petrozavodsk, pp.24-37.

УДК 512.544.43

**О ВОЗНИКНОВЕНИИ ГИГАНТСКОГО  $\varphi$ -ПРИМИТИВНОГО  
ИДЕАЛА СЛУЧАЙНОГО АВТОМОРФИЗМА  $\varphi$   
КОНЕЧНОЙ БУЛЕВОЙ РЕШЕТКИ**

Е. В. ХВОРОСТЯНСКАЯ

В работе рассматривается множество  $A_{n,N}$  всех автоморфизмов  $\varphi$  булевой решетки длины  $n$  таких, что число всех  $\varphi$ -примитивных идеалов равно  $N$ , на котором задано равномерное распределение вероятностей. Доказано, что гигантский  $\varphi$ -примитивный идеал случайного автоморфизма  $\varphi \in A_{n,N}$  возникает только в случае, когда  $n, N \rightarrow \infty$  так, что  $N/\ln n \rightarrow 0$ .

Пусть  $L_n$  — булева решетка длины  $n$  [1]. Обозначим через  $\text{Aut}(L_n)$  множество всех автоморфизмов булевой решетки  $L_n$ . Идеал  $I$  решетки  $L_n$  будем называть  $\varphi$ -примитивным, если  $\varphi(I) = I$  и из  $\varphi(x) = x$ ,  $x \in I$ , следует, что либо  $x$  равно наименьшему элементу решетки  $L_n$ , либо  $x$  равно наибольшему элементу идеала  $I$ . Обозначим через  $A_{n,N}$  множество всех автоморфизмов  $\varphi$  булевой решетки  $L_n$  таких, что число всех  $\varphi$ -примитивных идеалов равно  $N$ . Зададим на  $\text{Aut}(L_n)$  и  $A_{n,N}$  равномерное распределение вероятностей.

В работе [1] получены предельные теоремы для числа  $\varphi$ -примитивных идеалов случайного автоморфизма  $\varphi \in \text{Aut}(L_n)$  при  $n \rightarrow \infty$  и для числа  $\varphi$ -примитивных идеалов заданной длины случайного автоморфизма  $\varphi \in A_{n,N}$  при  $n, N \rightarrow \infty$ . Предложение 1 [1] показывает, что данные числовые характеристики автоморфизма

определяют длину решетки  $\text{Fix}_{L_n}(\varphi)$  неподвижных элементов автоморфизма  $\varphi$ , число  $\varphi$ -примитивных конгруэнций решетки  $L_n$ , а также число  $\varphi$ -примитивных конгруэнций  $\theta$  решетки  $L_n$  с заданной длиной фактор - решетки  $L_n/\theta$ .

Будем говорить, что при действии случайного автоморфизма  $\varphi \in A_{n,N}$  возникает гигантский  $\varphi$ -примитивный идеал, если при  $n \rightarrow \infty$  длина максимального  $\varphi$ -примитивного идеала имеет порядок  $n$  с вероятностью, стремящейся к 1, а следующий по величине  $\varphi$ -примитивный идеал имеет бесконечно малую в сравнении с  $n$  длину с вероятностью, стремящейся к 1.

**Теорема 1.** *При действии случайного автоморфизма  $\varphi \in A_{n,N}$  возникает гигантский  $\varphi$ -примитивный идеал только в случае, когда  $n, N \rightarrow \infty$  так, что  $N/\ln n \rightarrow 0$ .*

**ДОКАЗАТЕЛЬСТВО.** В [1] доказано, что существуют взаимно однозначные соответствия между множеством  $\text{Aut}(L_n)$  и множеством всех подстановок  $n$ -элементного множества  $A_n$  всех атомов решетки  $L_n$ , а также между  $\varphi$ -примитивными идеалами решетки  $L_n$  и циклами подстановки  $\varphi$  на множестве  $A_n$ . В [2] установлено, что в случайной подстановке степени  $n$ , имеющей  $N$  циклов, возникает гигантский цикл (т.е. длина максимального цикла имеет порядок  $n$  с вероятностью, стремящейся к 1, а следующий по величине цикл имеет бесконечно малую в сравнении с  $n$  длину с вероятностью, стремящейся к 1 только в случае, когда  $n, N \rightarrow \infty$  так, что  $N/\ln n \rightarrow 0$ . Отсюда следует утверждение теоремы.

## Литература

1. Тарарин В. М., Черепанова Е. В. О действии случайного автоморфизма конечной булевой решетки // Труды ИПМИ КарНЦ РАН, 2005, с.139–147.
2. Казимиров Н. И. Возникновение гигантской компоненты в случайной подстановке с известным числом циклов // Дискретная математика, 2003, 15, №3, с.128–144.

УДК 517.9

**СХОДИМОСТЬ СЕТОЧНЫХ АППРОКСИМАЦИЙ РЕШЕНИЯ  
КРАЕВОЙ ЗАДАЧИ С НЕЛИНЕЙНЫМ ГРАНИЧНЫМИ  
УСЛОВИЕМ И ПОДВИЖНОЙ ГРАНИЦЕЙ**

И. А. ЧЕРНОВ

В статье рассматривается нелинейная краевая задача типа Стефана, моделирующая гидридообразование при постоянных условиях. Предложен численный метод, позволяющий строить сеточные аппроксимации порядка  $\sqrt{h}$ ,  $h$  – шаг сетки. Доказана сходимость аппроксимаций к обобщенному непрерывному решению уравнения, удовлетворяющему краевым условиям. Подвижная граница получается гладкой. Тем самым доказано существование обобщенного решения задачи.

1. ВВЕДЕНИЕ

Гидриды металлов весьма перспективны для задач хранения и транспортировки водорода. Кинетика образования и разложения гидридов интенсивно изучается [1, 2]. Соответствующие модели имеют форму краевых задач с подвижной границей раздела фаз и нелинейными граничными условиями (физико-химические процессы на поверхностях) [3]. Для таких задач открыты вопросы существования и единственности решения и наличия сходящихся численных методов.

В этой работе мы рассмотрим одну модель гидридообразования при постоянных условиях; для нее построим сеточный метод решения: алгоритм, позволяющий получить сеточную аппроксимацию на произвольно мелкой сетке; докажем компактность последовательности аппроксимаций в пространстве непрерывных

функций. Предел удовлетворяет уравнению в обобщенном (слабом) смысле и граничным условиям и является обобщенным решением задачи. Единственность и устойчивость здесь не изучаем.

## 2. ПОСТАНОВКА ЗАДАЧИ

Рассматриваем гидрирование частицы порошка металла. Давление водорода и температура постоянны. Частицу представляем в виде тонкой пластинки-параллелепипеда, процессами на «узких» гранях пренебрегаем. Тогда диффузия внутри частицы одномерна (по глубине  $x \in [-L, L]$ ),  $2L$  — толщина пластинки. Осевая симметрия позволяет рассматривать только  $x \in [0, L]$ .

Считаем, что на поверхности пластинки к началу эксперимента уже образовался слой гидрида (корка) толщины  $L - \rho_0$ . Из-за фазового перехода корка становится толще, ее толщина  $L - \rho(t)$ . Функция  $\rho(t)$  описывает положение межфазной границы. Время  $t_{\text{end}}$  соответствует полному гидрированию образца ( $\rho(t_{\text{end}}) = 0$ ).

В гидриде химически связан водород; его концентрацию примем за единицу. Считаем, что в металле растворяется очень мало водорода по сравнению с единицей. В гидриде может растворяться и диффундировать еще водород с концентрацией  $c(t, x) - 1$ . Коэффициент диффузии примем за единицу.

На поверхности  $x = L$  происходит поглощение водорода (при достаточном давлении). Обозначим плотность потока сорбции через  $P = \text{const}$  (зависимостью его от давления не интересуемся).

Обычно полагают, что объемная десорбция квадратична относительно концентрации  $c(t, L)$  (из-за образования молекул  $H_2$ ). Однако концентрация водорода в гидриде слагается из химически связанныго водорода и растворенного. Есть две точки зрения: 1) образовывать молекулы могут все атомы водорода (химические связи относительны); 2) только подвижные атомы образуют молекулы. Обобщим оба варианта такой моделью: считаем плотность потока десорбции равной  $b(c(t, L) - \tilde{c})^2$ ,  $b = \text{const}$ . Здесь  $\tilde{c} \in [0, 1]$  — это «связанная» концентрация, не способная к десорбции. При  $\tilde{c} = 0$  имеем первую точку зрения, при  $\tilde{c} = 1$  — вторую. Полный поток сорбции имеет плотность  $J = P - b(c(t, L) - \tilde{c})^2$ . Из соображений сохранения следует, что поглощенный водород уносится вглубь образца диффузией: получаем граничное условие

$$\partial_x c(t, L) = P - b(c(t, L) - \bar{c})^2.$$

Рассмотрим процессы на границе раздела фаз  $x = \rho(t)$ . Считаем, что растворенный водород у границы раздела фаз будет потрачен на формирование гидрида; стало быть, чем больше концентрация, тем быстрее движется граница. При единичной концентрации формирование гидрида невозможно (нет избыточного водорода). Расход водорода за время  $dt$  на единице площади:  $-c(t, \rho)d\rho$  с учетом того, что  $d\rho < 0$ . Этот водород извлекается из приграничной области толщины  $kdt$ , коэффициент  $k$  положим равным единице. При малых  $dt$  концентрацию растворенного водорода в этой области можно считать постоянной и равной  $c(t, \rho) - 1$ , тогда извлекаемое количество (на единицу площади контакта фаз) составляет  $(c(t, \rho) - 1)dt$ . Из закона сохранения:

$$c(t, \rho(t))\dot{\rho}(t) = -(c(t, \rho(t)) - 1).$$

С другой стороны, диффузионно-подвижный водород приносится к границе раздела фаз диффузией. Если градиент концентрации мал по сравнению с избытком водорода, водород будет потрачен на формирование гидрида, концентрация понизится и градиент возрастет; если градиент велик, приток водорода увеличит приграничную концентрацию. Итак, градиент и концентрация связаны у границы. Эти рассуждения позволяют считать, что расход водорода на формирование гидрида обусловлен также диффузионным притоком к границе раздела фаз. Аналогичные выкладки приводят к условию типа Стефана:

$$c(t, \rho(t))\dot{\rho}(t) = -\partial_x c(t, \rho(t)).$$

Уточним начальное распределение водорода в корке (толщиной  $L - \rho_0$ ). Считаем корку тонкой, поэтому допустим, что в процессе ее формирования диффузия успела сделать профиль концентрации стационарным. Из уравнения диффузии получаем следующее условие на стационарное распределение  $\varphi(x)$ :  $\partial_{xx}\varphi(x) = 0$ ; следовательно,  $\varphi(x) = Bx + A$ .

Краевая задача, таким образом, имеет вид

$$\partial_t c(t, x) = \partial_{xx} c(t, x), \quad x \in (\rho(t), L), \quad t \in (0, t_{\text{end}}), \quad (1)$$

$$\partial_x c(t, L) = P - b(c(t, L) - \tilde{c})^2, \quad P > b(1 - \tilde{c})^2, \quad 0 \leq \tilde{c} \leq 1 \quad (2)$$

$$\partial_x c(t, \rho(t)) = c(t, \rho(t)) - 1, \quad (3)$$

$$\dot{\rho} = c(t, \rho(t))^{-1} - 1, \quad \rho(0) = \rho_0, \quad \rho(t_{\text{end}}) = 0, \quad (4)$$

$$c(0, x) = Bx + A, \quad x \in [\rho_0, L], \quad 0 < \rho_0 < L. \quad (5)$$

Неравенство  $P > b(1 - \tilde{c})^2$  означает физическую возможность сорбции: давление должно быть достаточно высоким, чтобы сорбция превосходила десорбцию при минимально-возможной стехиометрической концентрации (принятой за единицу).

Константы  $A$  и  $B$  определяются из условий баланса потоков (2), (3) при  $t = 0$ . Из (3) получим  $A = 1 + B(1 - \rho_0)$ . Величина  $B \geq 0$  по смыслу; тогда начальное распределение  $A + Bx \geq 1$  (и потому положительно) для всех  $x \in [\rho_0, L]$ :

$$A + Bx = 1 + B(1 - \rho_0) + Bx \geq 1 + B(1 - \rho_0) + B\rho_0 = 1 + B \geq 1.$$

Рассмотрим (2) при  $t = 0$ :  $B = P - b(A + BL - \tilde{c})^2$ . Подставляя  $A$ , получим квадратное уравнение (обозначим  $\hat{c} = 1 - \tilde{c}$ )

$$b(L - \rho_0 + 1)^2 B^2 + (2b\hat{c}(1 + L - \rho_0) + 1)B - (P - b\hat{c}^2) = 0.$$

Это уравнение имеет вещественные ненулевые корни разных знаков (см. неравенство в (2)); выбираем положительный.

Мы выбрали коэффициенты диффузии и гидридообразования и стехиометрическую концентрацию равными единице. Это допустимо, так как они имеют разную размерность ( $\text{см}^2/\text{сек}$ ,  $\text{см}/\text{сек}$  и моль/ $\text{см}^3$ ), а в задаче изначально три независимые размерности: длина, время, количество вещества. Других возможностей нет, поэтому параметры  $L$ ,  $\rho_0$ ,  $\tilde{c}$ ,  $P$  и  $b$  нормировать нельзя.

### 3. Принцип максимума для краевой задачи

Пусть  $\rho(t)$  – невозрастающая функция, заданная на  $[0, T]$  и имеющая производную, непрерывную в  $[0, T]$ . Пусть  $\rho(0) = \rho_0$  и  $\rho(T) \geq 0$ . Обозначим  $\mathbb{N} = [0, T] \times [0, L]$  и рассмотрим множество  $Y_\rho(T) \subset \mathbb{N}$  такое, что если точка  $(t, x) \in Y_\rho(T)$ , то  $0 < t < T$  и  $\rho(t) < x < L$ ; также введем замыкание  $\bar{Y}_\rho(T)$  множества  $Y_\rho(T)$ . Множества  $\bar{Y}_\rho(T)$  и  $Y_\rho(T)$  полностью определяются функцией  $\rho(t)$ .

Пусть  $c(t, x)$  задана в  $\bar{Y}_\rho(T)$ , имеет в  $Y_\rho(T)$  две частные производные по  $x$  и одну — по  $t$ , причем они непрерывны в  $\bar{Y}_\rho(T)$ . Пусть  $c(t, x)$  удовлетворяет в  $Y_\rho(T)$  уравнению диффузии (1), при  $x = L$  удовлетворяет (2), при  $x = \rho(t)$  удовлетворяет (3) и при  $t = 0$  удовлетворяет начальному условию (5). Пусть также функции  $\rho(t)$  и  $c(t, x)$  удовлетворяют при  $t \in [0, T]$  уравнению (4). Тогда пару  $\rho(t)$ ,  $c(t, x)$  назовем классическим решением задачи (1)–(5) в  $\bar{Y}_\rho(T)$ . Далее — классическое решение.

**Утверждение 1.** Пусть классическое решение  $c(t, x)$  существует в  $\bar{Y}_\rho(T)$  и его максимум или минимум достигается в точке  $(t^*, x^*)$ . Тогда либо  $t^* = 0$ , либо  $x^* = \rho(t^*)$ , либо  $x^* = L$ .

**Доказательство.** Множество  $\bar{Y}_\rho(T)$  замкнуто и ограничено и потому компактно. Непрерывная функция  $c(t, x)$  достигает на  $\bar{Y}_\rho(T)$  экстремальных значений.

Далее, в силу (5) решение не постоянно в  $\bar{Y}_\rho(T)$  при любых  $T$ .

Предположим,  $(t^*, x^*) \in Y_\rho(T)$  (внутренняя точка). Рассмотрим прямоугольник  $[t^* - \varepsilon, t^* + \varepsilon] \times [x^* - \delta, x^* + \delta] \subset Y_\rho(T)$ . Внутри этого прямоугольника функция  $c(t, x)$  непрерывна и удовлетворяет уравнению диффузии. По принципу максимума для параболических уравнений, экстремальное значение достигается на границе прямоугольника, что противоречит предположению.

Пусть максимум достигнут при  $t^* = T$ ,  $\rho(T) < x^* < L$ . Рассмотрим прямоугольник  $[T - \varepsilon, T] \times [x - \delta, x + \delta]$ , целиком лежащий в  $Y_\rho(T)$ . Далее рассуждаем аналогично.

**Утверждение 2.** Пусть классическое решение  $c(t, x)$  существует в  $\bar{Y}_\rho(T)$ . Тогда  $1 \leq c(t, x) \leq \sqrt{P/b} + \tilde{c}$ .

**Доказательство.** Пусть минимум  $c(t^*, x^*) < 1$  при  $(t^*, x^*) \in \bar{Y}_\rho(T)$ . Из утверждения 1 следует, что достаточно проверить значения при  $t^* = 0$ ,  $x^* = L$  и  $x^* = \rho(t^*)$ .

Однако  $c(0, x) = A + Bx \geq A + B\rho_0 = 1 + B > 1$ .

При  $x^* = \rho(t^*)$  применяем (3):  $\partial_x c(t^*, \rho(t^*)) = c(t^*, \rho(t^*)) - 1 < 0$ . Но это означает, что  $c(t^*, \rho(t^*) + \varepsilon) < c(t^*, \rho(t^*))$  при малых  $\varepsilon$ , а это противоречит тому, что в  $(t^*, x^*)$  достигается минимум.

При  $x^* = L$  положительный минимум достигаться не может. В самом деле, минимум заведомо меньше  $\sqrt{P/b} + \tilde{c}$ . Применим (2):

$$\partial_x c(t^*, L) = P - b(c(t^*, L) - \tilde{c})^2 > 0,$$

а значит,  $c(t^*, L - \varepsilon) < c(t^*, L)$  при малых  $\varepsilon$ , что и означает отсутствие минимума в  $(t^*, L)$ . Следует исключить бессмысленную с физической точки зрения ситуацию, когда  $c(t^*, L) < 0$ . Но если это так, то найдется такая точка  $(t^{**}, x^{**})$ , что  $c(t^{**}, x^{**}) = 0$  и  $c(t, x) > 0$  при  $t < t^{**}$ . Во множестве  $\bar{Y}_\rho(t^{**})$  функция  $c(t, x)$  достигает положительного минимума, который меньше единицы, а невозможность этого мы показали.

Пусть теперь  $c(t^*, x^*) > \sqrt{P/b} + \tilde{c}$  при  $(t^*, x^*) \in \bar{Y}_\rho(T)$ . Тогда максимум функции  $c(t, x)$  в  $\bar{Y}_\rho(T)$  тоже больше этого числа. Достаточно проверить значения при  $t^* = 0$ ,  $x^* = L$  и  $x^* = \rho(t)$ .

При  $x^* = L$  применим (2):

$$\partial_x c(t^*, L) = P - b(c(t^*, L) - \tilde{c})^2 < 0,$$

а значит,  $c(t^*, L - \varepsilon) > c(t^*, L)$  при малых  $\varepsilon$ : противоречие.

Пусть  $t^* = 0$ ; тогда  $c(0, x) = A + Bx \leq A + BL$ , так что  $x^* = L$ . Так как (2) применимо при  $t = 0$ , имеем уже рассмотренный случай.

При  $x^* = \rho(t^*)$  применяем (3):

$$\partial_x c(t^*, \rho(t^*)) = c(t^*, \rho(t^*)) - 1 > 0.$$

Имеем противоречие:  $c(t^*, \rho(t^*) + \varepsilon) > c(t^*, \rho(t^*))$  при малых  $\varepsilon$ .

*Следствие 1.* Пусть классические решения  $c(t, x)$  и  $u(t, x)$  существуют в  $\bar{Y}_\rho(T)$ . Тогда  $c(t, x) = u(t, x)$  в  $\bar{Y}_\rho(T)$ .

**ДОКАЗАТЕЛЬСТВО.** Предположим противное. Рассмотрим функцию  $w(t, x) = c(t, x) - u(t, x)$  и пусть  $w(t, x) > 0$  в какой-нибудь точке. Очевидно,  $w(t, x)$  удовлетворяет (1) и потому подчиняется принципу максимума. Ясно, что максимум положителен.

Однако  $w(0, x) = 0$ , поэтому положительный максимум при  $t = 0$  не достигается.

При  $x^* = \rho(t^*)$  функция  $w(t, x)$  удовлетворяет уравнению

$$\partial_x w(t^*, \rho(t^*)) = w(t^*, \rho(t^*)).$$

Если в точке  $(t^*, \rho(t^*))$  максимум, то  $\partial_x w(t^*, \rho(t^*)) > 0$ , что противоречит наличию максимума.

Рассмотрим случай  $x^* = L$ . Тогда  $w(t, x)$  удовлетворяет

$$\partial_x w(t, L) = -bw(t, L)\left(c(t, L) + u(t, L) + 2\tilde{c}\right).$$

По утверждению 2 скобка положительна. Поэтому при достижении в точке  $(t^*, L)$  положительного максимума получаем отрицательную производную  $\partial_x w(t, L)$  и тем самым противоречие.

Итак, получаем, что положительный максимум достигаться в  $\bar{Y}_\rho(T)$  не может, а потому  $w(t, x) \equiv 0$ , что и требовалось.

Поскольку мы доказали, что решение не меньше единицы и потому положительно, заменим (с целью технического удобства) (2) тождественным с ним условием

$$\partial_x c(t, L) = P - \operatorname{sgn}(c(t, L)) b \left( c(t, L) - \hat{c} \right)^2. \quad (6)$$

#### 4. Сеточная аппроксимация

Разобъем отрезок  $[\rho_0, L]$  на  $M$  равных частей длины  $h$ . Далее под обозначением  $h \rightarrow 0$  будем понимать  $M \rightarrow \infty$ .

Введем обозначения  $I = [L/h]$  (целая часть),  $K = I - M$  и  $\delta_h = L - Ih$ . Зададим какую-нибудь последовательность  $K_n$  с условиями:  $K_0 = K$ , а для  $n \geq 1$  либо  $K_{n+1} = K_n$ , либо  $K_{n+1} = K_n + 1$ . Для сокращения записей введем обозначение  $k_n = K_n - n$ .

Введем произвольную сетку, неравномерную по времени, с шагами  $h$  и  $\tau_n$ ,  $0 \leq n \leq N$ ,  $0 \leq i \leq I$ . Число  $N$  определим как как наименьшее  $N$ , отвечающее условию  $t_N = \sum_0^N \tau_n \geq T$ , где  $T$  – заданное время. Узлы этой сетки обозначим  $(n, i)$ : это точки  $(t_n, x_i)$ , где  $x_i = \delta_h + ih$ ,  $t_0 = 0$ ,  $t_{n+1} = t_n + \tau_n$ ,  $n > 0$ .

Рассмотрим подмножество  $\bar{D}_N$  сетки, состоящее из узлов  $(n, i)$ , таких, что для каждого  $0 \leq n \leq N$ ,  $i = k_n, k_n + 1, \dots, I$ . Иными словами, на каждом последующем слое либо на один узел больше слева (если  $K_{n+1} = K_n$ ), либо столько же узлов (если  $K_{n+1} = K_n + 1$ ). Введем также «открытое» подмножество сетки  $D_N$  с узлами  $(n, i)$  такими, что для  $0 < n < N$   $i = k_n + 1, \dots, I - 1$ . Еще нам понадобится сеточное множество  $\tilde{D}_N$ , полученное из  $\bar{D}_N$  удалением узлов  $(0, K)$  и  $(0, I)$ .

Узлы  $(n, k_n)$  являются сеточным аналогом подвижной границы. Ее непрерывную аппроксимацию получим, проведя впоследствии линейную интерполяцию [4].

Пример такого множества на сетке изображен на рисунке. Черные кружки – это узлы, входящие в  $D_N$ , звездочки – узлы, входящие в  $\bar{D}_N$ , но не в  $D_N$ . Подчеркнуты узлы  $(0, K)$  и  $(0, I)$ .

$$\circ \circ \circ \circ \circ \star \bullet \bullet \bullet \star \quad n = 2$$

$$\circ \circ \circ \circ \circ \star \bullet \bullet \bullet \star \quad n = 1$$

$$\circ \circ \circ \circ \circ \circ \underline{\star} \star \underline{\star} \quad n = 0$$

Сеточный аналог функции  $f(t, x)$  в узле  $(n, i)$  обозначим  $f_n^i$ .

Введем аппроксимации производных:

$$\partial_t c \approx \partial_\tau c_n^i = \frac{c_n^i - c_{n-1}^i}{\tau_{n-1}}, \quad \partial_x c \approx \partial_h c_n^i = \frac{c_n^i - c_{n+1}^i}{h}.$$

Заметим, что  $\partial_{xx} c \approx \partial_{hh} c_n^i = \partial_h(\partial_h c_n^{i+1})$ .

Будем считать  $\rho$  заданной. Ее значения на сетке (времени)  $\rho_n$ . Заменим производные и функции сеточными аналогами в (1), (3), (5) и (6). Получим систему уравнений:

$$\partial_\tau c_n^i = \partial_{hh} c_n^i, \quad (n, i) \in D_N, \quad (7)$$

$$\partial_h c_n^I = P - \operatorname{sgn}(c_n^I)b \left( c_n^I - \bar{c} \right)^2, \quad 0 \leq n \leq N, \quad (8)$$

$$\partial_h c_n^{k_n+1} = c_n^{k_n} - 1, \quad 0 \leq n \leq N, \quad (9)$$

$$c_0^i = A + Bi h, \quad i = K \dots I. \quad (10)$$

Параметры заданы, шаги  $h$  и  $\tau_n$  тоже считаем известными. Все рассматриваемые сеточные функции считаем заданными на  $\bar{D}_N$ .

## 5. ПРИНЦИП МАКСИМУМА ДЛЯ СЕТОЧНОЙ ЗАДАЧИ

Здесь мы докажем ряд утверждений, аналогичных принципу максимума и его следствиям для непрерывной задачи.

**Утверждение 3.** Пусть сеточная функция  $c_n^i$  задана в  $\bar{D}_N$  и удовлетворяет системе (7) в  $D_N$ , причем максимум или минимум (не обязательно строгие) достигается в узле  $(n^*, i^*)$ . Тогда либо  $n^* = 0$ , либо  $i^* = I$ , либо  $i^* = k_{n^*}$ , либо  $c_n^i = \text{const}$  в  $\bar{D}_{n^*}$ .

**ДОКАЗАТЕЛЬСТВО.** Пусть максимум (возможно, нестрогий) достигнут во внутреннем узле  $(n^*, i^*)$ ,  $n^* > 0$ ,  $K_{n^*} - n^* + 1 \leq i^* \leq \dots, I - 1$ . Тогда применимо уравнение (7). Левая часть его, очевидно, неотрицательна, а правая — неположительна. Это возможно, только если  $\partial_\tau c_{n^*}^{i^*} = \partial_{hh} c_{n^*}^{i^*} = 0$ . Следовательно,  $c_{n^*}^{i^* \pm 1} = c_{n^*}^{i^*}$  (поскольку это — максимум) и  $c_{n^*-1}^{i^*} = c_{n^*}^{i^*}$ . Применим рассуждение к узлам  $(n^*, i^* \pm 1)$ . Получим, что  $c_{n^*}^{i^* \pm 2} = c_{n^*}^{i^* \pm 1} = c_{n^*}^{i^*}$ . Продолжая рассуждения, получим  $c_{n^*}^i = c_{n^*}^{i^*}$  при  $i = K_{n^*} - n^* \dots I$  (функция на всем слое константа). Так как  $c_{n^*-1}^{i^*} = c_{n^*}^{i^*}$ , то максимум достигнут

также на слое  $n^* - 1$ . Аналогично получаем, что  $c_{n^*-2}^{i^*} = c_{n^*}^{i^*}$  и на всем узле слоя  $n^* - 1$  функция константа. Продолжая рассуждение, получим, что  $c_n^i = \text{const}$  в  $\bar{D}_{n^*}$ .

Для минимума доказательство аналогично.

*Следствие 2.* Пусть сеточная функция  $c_n^i$  задана в  $\bar{D}_N$  и удовлетворяет системе (7)–(10). Тогда если ее максимум или минимум (в том числе нестрогие) достигаются в узлах  $(n^*, i^*)$ , то либо  $n^* = 0$ , либо  $i^* = I$ , либо  $i^* = k_{n^*}$ .

*Доказательство.* В силу (10) функция  $c_n^i$  не константа при  $n = 0$ , и потому не константа в  $\bar{D}_n$  при любом  $n$ . См. утверждение 3.

**Утверждение 4.** Пусть сеточная функция  $c_n^i$  удовлетворяет системе (7)–(10) в  $\bar{D}_N$ . Тогда  $1 < c_n^i < \sqrt{P/b} + \tilde{c}$  в  $\bar{D}_N$ .

*Доказательство.* Достаточно доказать это неравенство для экстремальных значений. При  $n = 0$  оно верно: максимум  $c_0^i$  достигается при  $i = I$ ; если он равен или превосходит  $\sqrt{P/b} + \tilde{c}$ , то правая часть (8) неположительна, но левая часть (8) при  $n = 0$  равна  $B > 0$ . Минимум  $c_0^i$  достигается при  $i = K$  и больше 1.

Пусть максимум (он заведомо больше единицы) достигнут на левом краю, то есть, в узле  $(n^*, k_{n^*})$ . Из уравнения (9) имеем  $\partial_h c_{n^*}^{k_{n^*}+1} > 0$ , то есть  $c_{n^*}^{k_{n^*}+1} > c_{n^*}^{k_{n^*}}$ : противоречие.

Из (9) получаем, что если минимум (меньший единицы) достигнут в узле  $(n^*, k_{n^*})$ , то  $\partial_h c_{n^*}^{k_{n^*}+1} < 0$ : противоречие.

Осталось проверить правый край. Если максимум достигнут в узле  $(n^*, I)$ , то из (8) получаем, что при  $c_n^I > \sqrt{P/b} + \tilde{c}$  правая часть отрицательна, а значит,  $\partial_h c_n^I < 0$ . Опять получили противоречие.

Минимум заведомо меньше  $\sqrt{P/b} + \tilde{c}$ . Пусть он достигнут в узле  $(n^*, I)$ . Из уравнения (8) имеем  $\partial_h c_{n^*}^I > 0$ : противоречие. Здесь пригодился вспомогательный множитель в (8).

Пусть максимум  $c_{n^*}^I = \sqrt{P/b} + \tilde{c}$ . Тогда из (8) имеем  $\partial_h c_{n^*}^I = 0$ , то есть,  $c_{n^*}^I = c_{n^*}^{I-1}$ . Так как  $c_{n^*}^I$  — максимум, то максимум достигается во внутреннем узле  $= c_{n^*}^{I-1}$ , что по следствию 2 невозможно. Аналогично рассуждение для минимума  $c_{n^*}^{k_{n^*}}$ .

Так как сеточная функция  $c_n^i > 1$  в  $\bar{D}_N$ , вспомогательный множитель с  $\text{sgn}$  в (8) тождественно равен единице.

*Следствие 3.* Пусть  $c_n^i$  и  $w_n^i$  удовлетворяют системе (7)–(10) в  $\bar{D}_N$ . Тогда  $c_n^i = w_n^i$  в  $\bar{D}_N$ . (Решение единственно).

**ДОКАЗАТЕЛЬСТВО.** Доказательство опирается на доказанные утверждения и аналогично доказательству следствия 1.

Введем подмножество  $\bar{D}'_N$  сетки, состоящее из узлов  $(n, i)$ , таких, что для каждого  $0 \leq n \leq N$   $i = k_n + 1, k_n + 2, \dots, I$ . То есть, на один узел слева меньше, чем в  $\bar{D}_N$ . Введем также «открытое» подмножество сетки  $D'_N$  с узлами  $(n, i)$  такими, что для каждого  $0 < n < N$   $i = K_n - n + 2, \dots, I - 1$ . Множество  $\bar{D}'_N$  получим из  $\bar{D}'_N$  удалением узлов  $(0, K + 1)$  и  $(0, I)$ .

**Замечание 1.** Пусть  $c_n^i$  удовлетворяет системе (7)–(10) в  $\bar{D}_N$ . Сеточная производная  $\partial_h c_n^i$  определена (как сеточная функция) на  $\bar{D}'_N$  и удовлетворяет в  $D'_N$  системе (7). Поэтому она попадает в условия утверждения 3 (принцип максимума).

**Утверждение 5.** Пусть  $c_n^i$  удовлетворяет системе (7)–(10) в  $\bar{D}_N$ . Сеточная производная  $\partial_h c_n^i$  неотрицательна и ограничена сверху в  $\bar{D}'_N$ , причем оценка не зависит от  $h$ .

**ДОКАЗАТЕЛЬСТВО.** Так как максимум и минимум  $\partial_h c_n^i$  достигаются на «границе»  $D'_N$ , то достаточно проверить ограниченность значений на ней, то есть в узлах  $(0, i)$ ,  $(n, I)$  и  $(n, k_n + 1)$ .

В силу (10)  $\partial_h c_0^i = B > 0$  для всех  $i$ .

В силу (8) и утверждения 4 имеем ( $\hat{c} = 1 - \tilde{c}$ ):

$$\partial_h c_n^I = P - \operatorname{sgn}(c_n^I)b(c_n^I - \tilde{c})^2 = P - b(c_n^I - \tilde{c})^2 \in (0, P - b\hat{c}^2).$$

В силу (9) и утверждения 4 имеем:

$$\partial_h c_n^{k_n+1} = c_n^{k_n} - 1 \in (0, \sqrt{P/b} - \hat{c}).$$

Таким образом, имеем оценку для производной:

$$0 < \partial_h c_n^i < Z = \max(B, P - b\hat{c}^2, \sqrt{P/b} - \hat{c}), \quad (n, i) \in \bar{D}'_N.$$

Введем подмножество  $\bar{D}''_N$  сеточного множества  $\bar{D}_N$  так: узел  $(n, i) \in \bar{D}_N$  включаем в  $\bar{D}''_N$ , только если узел  $(n - 1, i)$  входит в сетку и принадлежит  $\bar{D}_N$ . Легко видеть, что в  $\bar{D}''_N$  индекс  $n > 0$  (слой  $n = 0$  не входит), а также не входят узлы с  $i = k_n =$

$K_n - n$  если  $K_n = K_{n-1}$ . Поэтому это множество порождается из сетки какой-то последовательностью  $K''_n$ ,  $n \geq 1$ . Определим также множество  $D''_N$  аналогично  $D_N$ .

**Замечание 2.** Пусть  $c_n^i$  удовлетворяет системе (7)–(10) в  $\bar{D}_N$ . Тогда сеточная производная  $\partial_\tau c_n^i$  определена (как сеточная функция) на  $\bar{D}_N''$  и удовлетворяет в  $\bar{D}_N''$  системе (7). Поэтому она сама попадает в условия утверждения 3 (принцип максимума).

**Утверждение 6.** Пусть  $c_n^i$  удовлетворяет системе (7)–(10) в  $\bar{D}_N$ . Сеточная производная  $\partial_\tau c_n^i$  неотрицательна и ограничена сверху, причем оценка не зависит от  $h$ .

**Доказательство.** Так как максимум и минимум  $\partial_\tau c_n^i$  достигаются на «границе»  $D_N''$ , то достаточно проверить ограниченность значений на ней, то есть в узлах  $(1, i)$ ,  $(n, I)$  и  $(n, k''_n)$ .

Рассмотрим два узла  $(n+1, I)$  и  $(n, I)$ ,  $n \geq 0$ . В силу (8) имеем

$$\begin{cases} \frac{c_{n+1}^I - c_{n+1}^{I-1}}{h} = P - b(c_{n+1}^I - \tilde{c})^2, \\ \frac{c_n^I - c_n^{I-1}}{h} = P - b(c_n^I - \tilde{c})^2. \end{cases}$$

Вычтем из первого второе, разделим на  $\tau_n$ , умножим на  $h$ :

$$\partial_\tau c_{n+1}^I - \partial_\tau c_{n+1}^{I-1} = -hb(c_{n+1}^I + c_n^I + 2\tilde{c})\partial_\tau c_{n+1}^I.$$

Скобка в правой части больше двух и ограничена; поэтому

$$\partial_\tau c_{n+1}^I (1 + hb(c_{n+1}^I + c_n^I + 2\tilde{c})) = \partial_\tau c_{n+1}^{I-1}.$$

Это означает, что  $|\partial_\tau c_{n+1}^I| < |\partial_\tau c_{n+1}^{I-1}|$  и знаки одинаковые.

Отсюда мы делаем вывод, что положительный максимум и отрицательный минимум функции  $\partial_\tau c_n^i$  не могут достигаться при  $i = I$ : если  $\partial_\tau c_n^I$  — положительный максимум, то  $\partial_\tau c_n^{I-1} > \partial_\tau c_n^I$  — противоречие. Аналогично с отрицательным минимумом.

Теперь рассмотрим два узла  $(n+1, k''_{n+1})$  и  $(n, k''_n)$ ,  $n \geq 0$ . Сначала изучим случай  $K''_{n+1} = K''_n + 1$ . Тогда в силу (9)

$$\begin{cases} \frac{c_{n+1}^{k''_{n+1}+1} - c_{n+1}^{k''_n}}{h} = c_{n+1}^{k''_n} - 1, \\ \frac{c_n^{k''_n+1} - c_n^{k''_n}}{h} = c_n^{k''_n} - 1. \end{cases}$$

Вычтем из первого второе, разделим на  $\tau_n$ , умножим на  $h$ :

$$\partial_\tau c_{n+1}^{k''+1} - \partial_\tau c_{n+1}^{k''} = h \partial_\tau c_{n+1}^{k''}.$$

Поэтому  $(1+h)\partial_\tau c_{n+1}^{k''-n} = \partial_\tau c_{n+1}^{k''-n+1}$ , то есть  $|\partial_\tau c_{n+1}^{k''-n}| < |\partial_\tau c_{n+1}^{k''-n+1}|$  и знаки одинаковые. Аналогичное приведенному выше рассуждение показывает, что положительный максимум и отрицательный минимум функции  $\partial_\tau c_n^i$  не могут достигаться в узле  $(n, k_n)$ , если  $K_n'' = K_{n-1}'' + 1$ .

Теперь изучим случай, когда  $K_{n+1}'' = K_n''$ . В силу (9)

$$\begin{cases} \frac{c_{n+1}^{k''} - c_{n+1}^{k''-1}}{h} = c_{n+1}^{k''-1} - 1, \\ \frac{c_n^{k''+1} - c_n^{k''}}{h} = c_n^{k''} - 1. \end{cases}$$

Вычтем из второго первое и умножим на  $h$ :

$$(c_{n+1}^{k''-1} - c_n^{k''}) + (c_n^{k''+1} - c_{n+1}^{k''}) = h(c_n^{k''} - c_{n+1}^{k''-1}).$$

Упростим:

$$(1+h)(c_{n+1}^{k''-1} - c_n^{k''}) = -(c_n^{k''+1} - c_{n+1}^{k''}).$$

Рассмотрим отдельно вторую скобку:

$$\begin{aligned} c_{n+1}^{k''-1} - c_n^{k''} &= c_{n+1}^{k''-1} - c_{n+1}^{k''} + c_{n+1}^{k''} - c_n^{k''} = \\ &= -\partial_h c_{n+1}^{k''} h + \partial_\tau c_{n+1}^{k''} \tau_n. \end{aligned}$$

Аналогично с третьей скобкой:

$$c_n^{k''+1} - c_{n+1}^{k''} = \partial_h c_{n+1}^{k''+1} h - \partial_\tau c_{n+1}^{k''+1} \tau_n.$$

Подставляем полученные выражения и преобразуем:

$$(1+h)\partial_\tau c_{n+1}^{k''} = \partial_\tau c_{n+1}^{k''+1} + \frac{h}{\tau_n} (\partial_h c_{n+1}^{k''} - \partial_h c_{n+1}^{k''+1}) + \frac{h^2}{\tau_n} \partial_h c_{n+1}^{k''}.$$

Скобка в правой части равна  $-\partial_h c_{n+1}^{k''} h = -\partial_\tau c_{n+1}^{k''} h$  (применили уравнение (7)). Поэтому имеем

$$\left(1 + h + \frac{h^2}{\tau_n}\right) \partial_\tau c_{n+1}^{k''} = \partial_\tau c_{n+1}^{k''+1} + \frac{h^2}{\tau_n} \partial_h c_{n+1}^{k''} \quad (11)$$

Отрицательный минимум достигнут быть не может, потому что

$$\left(1 + h + \frac{h^2}{\tau_n}\right) \partial_\tau c_{n+1}^{k''_n} \geq \partial_\tau c_{n+1}^{k''_n+1}.$$

Если если  $\partial_\tau c_{n+1}^{k''_n} < 0$  — минимум, то  $\partial_\tau c_{n+1}^{k''_n+1} < 0$  и

$$\left(1 + h + \frac{h^2}{\tau_n}\right) |\partial_\tau c_{n+1}^{k''_n}| \leq |\partial_\tau c_{n+1}^{k''_n+1}|.$$

Значит,  $|\partial_\tau c_{n+1}^{k''_n}| \leq |\partial_\tau c_{n+1}^{k''_n+1}|$  и потому  $\partial_\tau c_{n+1}^{k''_n} \geq \partial_\tau c_{n+1}^{k''_n+1}$ .

Продолжаем рассуждение. Перепишем (11) в виде

$$(1 + h) \partial_\tau c_{n+1}^{k''_n} = \partial_\tau c_{n+1}^{k''_n+1} + \frac{h^2}{\tau_n} (\partial_h c_{n+1}^{k''_n} - \partial_\tau c_{n+1}^{k''_n}),$$

и тогда в силу ограниченности  $0 \leq \partial_h c_n^i \leq Z$

$$(1 + h) \partial_\tau c_{n+1}^{k''_n} \leq \partial_\tau c_{n+1}^{k''_n+1} + \frac{h^2}{\tau_n} (Z - \partial_\tau c_{n+1}^{k''_n}).$$

Отсюда следует, что если  $\partial_\tau c_{n+1}^{k''_n}$  достигает положительного максимума, то не большего  $Z$ . В противном случае первая скобка больше единицы, а последняя скобка отрицательна, поэтому  $\partial_\tau c_{n+1}^{k''_n} \leq \partial_\tau c_{n+1}^{k''_n+1}$ , что противоречит наличию максимума.

Осталось изучить случай  $n = 1$ . Напомним, мы доказываем ограниченность сеточной функции  $\partial_\tau c_n^i$ . Границные узлы  $i = I$  и  $i = k''_1 = K$  уже изучены. Введем обозначение  $\varepsilon^i = c_1^i - c_0^i$  и применим систему (7) при  $n = 1, K - 1 \leq i \leq I$ :

$$\partial_\tau c_1^i = \frac{c_1^{i-1} - 2c_1^i + c_1^{i+1}}{h^2} = \frac{c_0^{i-1} - 2c_0^i + c_0^{i+1}}{h^2} + \frac{\varepsilon^{i-1} - 2\varepsilon^i + \varepsilon^{i+1}}{h^2}.$$

Но  $c_0^{i-1} - 2c_0^i + c_0^{i+1} = c_0^i - Bh - 2c_0^i + c_0^i + Bh = 0$  из-за вида начальных условий (10). Кроме того,  $\partial_\tau c_1^i = \varepsilon_i / \tau_0$ . Окончательно имеем

$$h^2 \varepsilon_i = \tau_0 (\varepsilon^{i-1} - 2\varepsilon^i + \varepsilon^{i+1}).$$

Пусть в узле  $(1, i^*)$  функция  $\partial_\tau c_n^i$  достигла положительного максимума; тогда  $\varepsilon^{i^*}$  тоже достигает положительного максимума. При этом левая часть положительна, а правая отрицательна — противоречие. Аналогично покажем отсутствие отрицательного минимума при  $n = 1$ .

Подведем итоги. Минимальное и максимальное значение сеточной функции  $\partial_\tau c_n^i$  в  $\bar{D}_N^n$  достигается в узлах  $(n^*, i^*)$  при  $n^* = 1$ ,

$i^* = I$  или  $i^* = K''_{n^*} - n^*$ . Однако отрицательный минимум достигаться не может ни в одном случае; значит,  $\partial_\tau c_n^i \geq 0$  в  $\bar{D}_N''$ . Отсюда следует, что максимум  $\partial_\tau c_n^i$  в  $\bar{D}_N''$  положителен (или ноль, если  $\partial_\tau c_n^i \equiv 0$ ). Он не может быть достигнут при  $n^* = 1$  и при  $i^* = I$ , а при  $i^* = K_{n^*} - n^*$  не превосходит  $Z$ , которая не зависит от  $h$ . Поэтому  $0 \leq \partial_\tau c_n^i \leq Z$  в  $\bar{D}_N''$ , что и требовалось.

Подчеркнем, что обе сеточные производные функции  $c_n^i$  неотрицательны и ограничены константой  $Z$ , не зависящей от  $h$ .

*Следствие 4.* Покажем, что сеточная производная на правом краю  $\partial_h c_n^I$  решения  $c_n^i$  системы (\*) отделена от нуля; то есть, существует такое  $\gamma > 0$ , что  $\partial_h c_n^I \geq \gamma$  при  $n = 0, 1, \dots, N$ .

**Доказательство.** Предположим, что  $\partial_h c_{n^*}^I < \gamma$  при каком-то  $n^*$ . Теорема 9 утверждает, что сеточная производная  $\partial_\tau c_n^i$  неотрицательна в тех узлах сетки, где она определена. Следовательно, в силу (7), неотрицательна и  $\partial_{hh} c_n^i$ ; отсюда получаем, что сеточная функция  $\partial_h c_n^i$  не убывает по  $i$  при заданном  $n$ . Значит, при  $n = n^*$  сеточная функция  $\partial_h c_{n^*}^i < \gamma$  при  $i = k_{n^*}, \dots, I$ .

В силу (9)  $c_{n^*}^{k_{n^*}} < 1 + \gamma$ . Тогда

$$\begin{aligned} c_{n^*}^I &< 1 + \gamma + \sum_{i=k_{n^*}+1}^I \partial_h c_i^{n^*} h < 1 + \gamma + \gamma h(I - k_{n^*}) \leq \\ &\leq 1 + \gamma + \gamma h I \leq 1 + \gamma + \gamma(L - \delta_h) \leq 1 + \gamma(L + 1). \end{aligned}$$

Но в силу (8)  $c_{n^*}^I > \sqrt{(P - \gamma)/b} + \tilde{c}$ . Поэтому  $c_{n^*}^I$  подчинена двойному неравенству. Чтобы получить противоречие, нужно сделать левую часть не меньше правой:  $\tilde{c} + \sqrt{(P - \gamma)/b} \geq 1 + \gamma(L + 1)$ . Выполняя элементарные преобразования, получаем

$$b(L + 1)^2 \gamma^2 + (2b\hat{c}(L + 1) + 1)\gamma - (P - b\hat{c}^2) \leq 0.$$

Этот квадратный трехчлен имеет два корня разных знаков, положительный корень  $\bar{\gamma}$ . Решение неравенства  $0 < \gamma < \bar{\gamma}$ . При таких  $\gamma$  имеем противоречие, следовательно, предположение  $\partial_h c_{n^*}^I < \gamma$  неверно; таким образом,  $\partial_h c_n^I \geq \bar{\gamma}$  при всех  $n$ .

## 6. ПОДБОР ШАГА ПО ВРЕМЕНИ

До сих пор мы считали сетку заданной, то есть наряду с шагом  $h$  были известны и шаги  $\tau_n$ . Заданной считали также и последовательность  $K_n$ , которая определяет сеточную область определения  $\bar{D}_N$  задачи. Следует дополнить систему (7)–(10), чтобы находить из нее, кроме  $c_n^k$ , также  $\tau_n$  и  $K_n$ . Для этого рассмотрим сеточный аналог уравнения Стефана (4):

$$\partial_\tau \rho_{n+1} = (\rho_{n+1} - \rho_n)/\tau_n = (c_n^{k_n})^{-1} - 1. \quad (12)$$

Мы знаем, что  $c_n^{k_n} > 1$  и тем самым смещение подвижной границы  $\rho_{n+1} - \rho_n < 0$ . Чтобы это смещение при переходе от слоя к слою было в пределах точности равно по модулю шагу  $h$ , требуется специальный выбор шага  $\tau_n$ :

$$\tau_n = c_n^{k_n} (c_n^{k_n} - 1)^{-1} h.$$

Однако не исключено, что  $c_n^{k_n}$  близко к единице, и тогда  $\tau_n$  может быть недопустимо велико. Мы должны гарантировать, что  $\tau_n \rightarrow 0$  при  $h \rightarrow 0$ , иначе нет аппроксимации.

Уравнение (4) связывает скорость движения границы с отклонением  $c_n^{k_n}$  от единицы. Поэтому будем считать, что при «малых»  $c_n^{k_n}$  граница неподвижна. «Малые»  $c_n^{k_n}$  определим так:

$$c_n^{k_n} - 1 < B\sqrt{h}. \quad (13)$$

Положительность левой части гарантируется доказанными утверждениями. На шаг  $h$  наложим условие

$$h \leq 1. \quad (14)$$

Его смысл в том, чтобы начальное значение на подвижной границе  $c_0^K = 1 + B$  не оказалось «мало».

Таким образом, шаг  $\tau_n$  выберем так:

$$\tau_n = \begin{cases} \frac{c_n^{k_n}}{c_n^{k_n} - 1} h, & c_n^{k_n} - 1 \geq B\sqrt{h} \\ \frac{1 + B\sqrt{h}}{B} \sqrt{h} = \bar{\tau}, & \text{иначе} \end{cases}, \quad 0 \leq n \leq N - 1. \quad (15)$$

Максимальный шаг обозначен  $\bar{\tau} \leq \sqrt{h}(1 + B^{-1}) = O(\sqrt{h})$ . В первом случае сеточная подвижная граница смещается на один узел, а

во втором остается неподвижной; в первом случае мы добавляем к новому слою один узел слева, а во втором — нет. Поэтому

$$K_{n+1} = \begin{cases} K_n, & c_n^{k_n} - 1 \geq B\sqrt{h} \\ K_n + 1, & c_n^{k_n} - 1 < B\sqrt{h} \end{cases}, \quad 0 \leq n \leq N-1. \quad (16)$$

Добавим (15) и (16) к системе (7)–(10). Получим систему (\*).

*Замечание 3.* Пусть  $c_n^i$  удовлетворяет системе (\*) в  $\bar{D}_1$ . Заметим, что  $K_1 = K_0 = K$  в силу (14). Поэтому на слое  $n = 1$  на один узел больше, чем на слое  $n = 0$ . В самом деле,  $\bar{D}_1$  состоит из узлов  $(n, i)$ ,  $0 \leq n \leq 1$ ,  $K - n \leq i \leq I$ . Пусть сеточная производная  $\partial_h c_n^i$  равна константе в  $\bar{D}'_1$ . Тогда в силу (10)  $\partial_h c_n^i = B$  в  $\bar{D}'_1$ . Далее, в силу (9)  $c_0^K = c_1^{K-1} = 1 + B$  и  $c_1^K = c_1^{K-1} + \partial_h c_1^K h = 1 + B + Bh$ . Аналогично,  $c_1^{K+k} = c_1^K + \partial_h c_1^K kh = 1 + B + Bh + Bkh = 1 + B + B(k+1)h$ . Поэтому  $c_1^N = 1 + B + B(N-K+1)h$ . Рассуждая так же, получим  $c_0^N = 1 + B + B(N-K)h \neq c_1^N$ . В силу (6), однако,

$$c_0^N = c_1^N = \sqrt{(P-B)/b} + \tilde{c}.$$

Полученное противоречие означает, что  $\partial_h c_n^i$  не может быть константой в  $\bar{D}'_1$ , а значит, и в  $\bar{D}'_N$  при  $N \geq 1$ .

## 7. РЕШЕНИЕ СИСТЕМЫ

Докажем, что система (\*) (напомним, что это система (7)–(10) и (15), (16)) имеет решение, указав алгоритм его построения. Входные данные:  $P$ ,  $b$ ,  $\tilde{c}$ ,  $T$ ,  $L$ ,  $\rho_0$  и  $M$ . Решение — это набор  $\tau_n$  при  $0 \leq n \leq N-1$ ,  $K_n$  при  $0 \leq n \leq N$ ,  $c_n^i$  при  $0 \leq n \leq N$ ,  $k_n \leq i \leq I$ ; при этом однозначно определены  $h$  (оно удовлетворяет (14)),  $I$ ,  $K = K_0$  и  $N$  (см. стр. 103). Напомним, что  $k_n = K_n - n$ .

Алгоритм:

- (1) Пусть  $n = 0$ . Положим  $K_0 = K$ . При  $K_0 \leq i \leq I$  определяем  $c_0^i$  из (10).
- (2) Определяем  $\tau_n$  из (15). Очевидно,  $h < \tau_n \leq \bar{\tau}$ .
- (3) Определяем  $K_{n+1}$  из (16). Если  $K_{n+1} = n+1$  (т.е.  $k_{n+1} = 0$ ), задача решена (граница достигла нуля).
- (4) Увеличим  $n$  на единицу. Если  $n = N$ , задача решена.
- (5) Выразим из линейного уравнения (9)  $c_n^{k_n}$  через  $c_n^{k_n+1}$ :

$$c_n^{k_n} = \frac{c_n^{k_n+1} + h}{1+h} = \alpha_{k_n} c_n^{k_n+1} + \beta_{k_n}.$$

Коэффициенты линейной связи выпишем отдельно:

$$\alpha_{k_n} = (1 + h)^{-1}, \quad \beta_{k_n} = h(1 + h)^{-1}. \quad (17)$$

- (6) Выразим все  $c_n^i$ ,  $i = k_n + 1, \dots, I - 1$  через  $c_n^{i+1}$  линейно с неопределенными коэффициентами (метод прогонки):

$$c_n^i = \alpha_i c_n^{i+1} + \beta_i. \quad (18)$$

- (7) Подставим (18) в (7) вместо  $c_n^{i-1}$ . Получим ( $S_n = h^2/\tau_{n-1}$ )

$$\alpha_i = (2 + S_n - \alpha_{i-1})^{-1}, \quad \beta_i = \alpha_i(S_n c_n^{i-1} + \beta_{i-1}). \quad (19)$$

Это — рекуррентная последовательность с начальными данными (17),  $S_n$  известно;  $\alpha_i$  и  $\beta_i$  при  $k_n \leq i \leq I - 1$  определяются однозначно.

*Замечание 4.* Докажем по индукции, что  $\alpha_i \in (0, 1)$ ,  $\beta_i > 0$  при  $k_n \leq i \leq I - 1$ . Для  $i = k_n$  это так, см. (17). Если это так для  $i - 1$ , то из (19) это верно и для  $i$ .

- (8) Подставим (18) при  $i = I - 1$  в (8):

$$bh(c_n^I)^2 + c_n^I(1 - \alpha_{I-1} - 2hb\tilde{c}) - ((P - b\tilde{c}^2)h + \beta_{I-1}) = 0. \quad (20)$$

Из замечания 4 следует, что свободный член этого уравнения меньше нуля: корни вещественны и разных знаков.

Выбираем положительный:  $c_n^I$  определяется однозначно.

- (9) Пользуясь связью (18) и формулами (19) и (17), последовательно определяем  $c_n^i$  для  $i = I - 1, I - 2, \dots, k_n$

- (10) Идем на шаг 2.

Очевидно, алгоритм завершается за конечное число итераций: либо на шаге 3 (при достижении границей нуля), либо на шаге 4 (по истечении расчетного времени).

**Утверждение 7.** Пусть  $c_n^i$ ,  $\tau_n$  и  $K_n$  есть решение системы (\*) и  $h$  достаточно мало. Тогда существуют времена  $T_-$  и  $T_+$ , не зависящие от  $h$ , такие, что при  $T = T_-$  алгоритм завершится на шаге 4 (граница не достигнет нуля за данное время), а при  $T = T_+$  алгоритм завершится на шаге 3 (граница достигнет нуля до истечения данного времени).

**ДОКАЗАТЕЛЬСТВО.** Сначала установим оценку для времени  $T_-$ . За шаг по времени граница либо смещается на один узел влево, либо остается в прежнем узле. Изначально она занимала узел с

индексом  $K$ , следовательно, нужно не менее  $K$  шагов по времени для достижения нуля. Шаги по времени  $\tau_n > h$ , поэтому

$$T_- \geq Kh = (I - M)h = Ih - M(L - \rho_0)/M = \rho_0 - \delta_h \geq \rho_0 - h.$$

Если  $h$  достаточно мало, то есть  $h \leq \bar{h} < \rho_0$  (очевидное требование), то  $T_- \geq \rho_0 - \bar{h}$ . Оценка грубая: она означает, что скорость границы равна  $-1$ ; тогда из (12) следует, что  $c_n^{k_n} = \infty$ .

Теперь оценим  $T_+$ . Для этого рассмотрим сеточный аналог балансового уравнения (закона сохранения).

Пусть  $K_{n+1} = K_n + 1$ , то есть, граница неподвижна при переходе от слоя  $n$  к слою  $n+1$ ; при этом  $\tau_n = \bar{\tau}$ . Преобразуем приращение количества вещества за время  $\tau_n$  (суммы — количества, «содержащиеся» в слоях):

$$\begin{aligned} \sum_{i=k_{n+1}}^{I-1} c_{n+1}^i h - \sum_{i=k_n}^{I-1} c_n^i h &= \sum_{i=k_n}^{I-1} \partial_\tau c_{n+1}^i \tau_n h \geq \sum_{i=k_n+1}^{I-1} \partial_{hh} c_{n+1}^i h \tau_n = \\ &= \sum_{i=k_n+1}^{I-1} (\partial_h c_{n+1}^{i+1} - \partial_h c_{n+1}^i) \tau_n = (\partial_h c_{n+1}^I - \partial_h c_{n+1}^{k_n+1}) \tau_n. \end{aligned}$$

Здесь используем определение сеточной производной  $\partial_\tau c_n^i$ , уравнение (7), определение сеточной производной  $\partial_{hh} c_n^i$  и положительность  $\partial_\tau c_{n+1}^{k_n}$  (утверждение 6). Теперь оценим производную:

$$\begin{aligned} -\partial_h c_{n+1}^{k_n+1} &= -c_{n+1}^{k_n} + 1 = -\partial_\tau c_{n+1}^{k_n} \tau_n - c_n^{k_n} + 1 = \\ &= -\partial_\tau c_{n+1}^{k_n} \bar{\tau} - \partial_h c_n^{k_n+1} \geq -d\sqrt{h} - B\sqrt{h} = -(B+d)\sqrt{h}. \end{aligned}$$

Последняя оценка получается из определения  $\bar{\tau}$  (15) и «малости»  $c_n^{k_n}$ . Константа  $d > 0$  от  $h$  не зависит. Продолжаем:

$$\sum_{i=k_{n+1}}^{I-1} c_{n+1}^i h - \sum_{i=k_n}^{I-1} c_n^i h \geq (\bar{\gamma} - (B+d)\sqrt{h}) \tau_n. \quad (21)$$

Применили следствие 4. При достаточно малом  $h$  последняя скобка положительна и даже больше какой-нибудь константы  $\Gamma > 0$ .

Пусть  $K_{n+1} = K_n$ , то есть, граница смещается на шаг при переходе от слоя  $n$  к слою  $n + 1$ . Тогда (аналогично):

$$\begin{aligned} \sum_{i=k_{n+1}}^{I-1} c_{n+1}^i h - \sum_{i=k_n}^{I-1} c_n^i h &= c_{n+1}^{k_n+1} h + \sum_{i=k_n}^{I-1} (c_{n+1}^i - c_n^i) h = \\ &= c_{n+1}^{k_n+1} h + (\partial_h c_{n+1}^I - \partial_h c_{n+1}^{k_n}) \tau_n = \partial_h c_{n+1}^I \tau_n \geq \bar{\gamma} \tau_n \geq \Gamma \tau_n. \end{aligned} \quad (22)$$

Применили (9) и (15). Теперь оценим приращение количества с самого начала (помним, что  $t_n = \sum \tau_m$ ):

$$\sum_{i=k_{n+1}}^{I-1} c_{n+1}^i h - \sum_{i=K}^{I-1} c_0^i h = \sum_{m=0}^n \left( \sum_{i=k_{m+1}}^{I-1} c_{m+1}^i h - \sum_{i=k_m}^{I-1} c_m^i h \right) \geq \Gamma t_n.$$

Здесь мы использовали оценки (21) и (22). Но, с другой стороны,

$$\sum_{i=k_{n+1}}^{I-1} c_{n+1}^i h - \sum_{i=K}^{I-1} c_0^i h \leq \sum_{i=0}^{I-1} h \left( \sqrt{P/b} + \tilde{c} \right) \leq \left( \sqrt{P/b} + \tilde{c} \right) L$$

(отбросили вторую сумму и учли, что  $c_n^i \leq c_n^I$ ). Итак, при

$$t_n > L \Gamma^{-1} \left( \sqrt{P/b} + \tilde{c} \right) = T_+(h)$$

противоречие: граница достигнет нуля до  $T_+(h)$  при условии, что  $h$  мало; чем меньше  $h$ , тем больше  $\Gamma$  и тем лучше оценка.

Выясним порядок аппроксимации положительного точного гладкого решения (предположим, что оно существует). Напомним, что  $\tau_n = O(\sqrt{h})$ , см. (15). Погрешность аппроксимации (7) уравнения диффузии (1) оценивается  $O(\max(\tau_n) + h) = O(\sqrt{h})$  [5]. Погрешность аппроксимаций (8) и (9) граничных условий получается  $O(h)$ . Начальное условие (10) выполняется точно. Аппроксимация (12) условия Стефана (4) оценивается  $O(\sqrt{h})$ . Итак, порядок аппроксимации построенной разностной схемы  $\sqrt{h}$ .

*Замечание 5.* Порядок аппроксимации нельзя поднять выше  $\sqrt{h}$ , просто выбрав условие малости (13) в форме  $c_n^{K_n-n} - 1 < Bh^{1-\varepsilon}$ . При этом условие (14) на  $h$  не изменится и из (15) получим

$$h < \tau_n \leq \bar{\tau} = (B^{-1} + h^{1-\varepsilon})h^\varepsilon \leq (1 + B^{-1})h^\varepsilon.$$

Порядок аппроксимации (1)–(3) станет  $\varepsilon$ , но для условия Стефана он ухудшится  $(1 - \varepsilon)$ . Его можно улучшить до единицы, если доказать, что  $c_n^{K_n-n} - 1$  отделена от нуля.

### 8. Сходимость аппроксимаций

При каждом  $h$  можно однозначно получить решение  $\tau_n$  при  $0 \leq n \leq N-1$ ,  $\bar{D}_N$ ,  $c_n^i$  в  $\bar{D}_N$ . Дальнейшие рассуждения используют критерий компактности Арцела (Арчела) [6]

Определим  $\rho_h(t)$  линейной интерполяцией точек  $(t_n, x_{k_n})$ . Функция  $\rho_h(t)$  имеет кусочно-постоянную производную между узлами; оценим ее. Ясно, что  $-h \leq x_{k_{n+1}} - x_{k_n} \leq 0$  и  $t_{n+1} - t_n = \tau_n > h$ . Поэтому  $-1 \leq \dot{\rho}_h(t) \leq 0$  между узлами при любом  $h$ .

Сама  $\rho_h(t) \in [0, \rho_0]$  при любых  $h$ . По утверждению 9 при достаточно малом расчетном времени  $T$  граница в ноль не обращается.

Рассмотрим какую-нибудь последовательность  $h_j \rightarrow 0$ . Последовательность функций  $\rho_{h_j}(t)$  на  $[0, T]$  является равномерно ограниченной и равностепенно непрерывной; следовательно, по теореме Арцела, из нее можно выделить равномерно сходящуюся подпоследовательность (возможно, не одну). Подмножество индексов опять обозначим  $j$ . Пусть предел  $\rho(t)$ . Ясно, что  $\rho(t) > 0$ ,  $\rho(0) = \rho_0$  и не возрастает. Это — непрерывная подвижная граница. Увеличивая  $T$  и при необходимости продлевая  $\rho_h(t)$  нулем по непрерывности, покажем, что  $\rho(t_{\text{end}}) = 0$  при каком-то  $t_{\text{end}} \leq L(\bar{\gamma})^{-1}(\sqrt{P/b} + \bar{c}) < \infty$  (см. утверждение 9).

Теперь построим функцию  $\psi_h(t, x)$  в  $\mathbb{N} = [0, T] \times [0, L]$  следующим образом. В узлах  $\bar{D}_N$  положим  $\psi_h(t_n, x_i) = c_n^i$ . В узлах на один дальше крайнего левого на каждом слое положим  $\psi_h(t_n, x_{k_n-1}) = c_n^{k_n} - \partial_h c_n^{k_n+1} h$  (по «непрерывности» производной). В остальных узлах  $(n, i)$  сетки (при  $0 \leq n \leq N$ ,  $0 \leq i < k_n - 1$ ) положим  $\psi_h(t_n, x_i) = \psi_h(t_n, x_{k_n-1})$  («константой по непрерывности»).

В точках  $(t, x)$  при  $t_n \leq t \leq t_{n+1}$ ,  $x_i \leq x \leq x_{i+1}$  определим  $\psi_h(t, x)$  линейной интерполяцией. Пусть  $\alpha, \beta \in [0, 1]$ . Введем обозначение  $u_\alpha^\beta = \psi_h(t_n + \alpha\tau_n, x_i + h\beta)$ . Положим ( $s = 0$  или 1)

$$u^s = u_0^s + \alpha(u_1^s - u_0^s), \quad u_s^\beta = u_s^0 + \beta(u_s^1 - u_s^0), \quad u_\alpha^\beta = u_\alpha^0 + \beta(u_\alpha^1 - u_\alpha^0).$$

Ясно, что  $u_\alpha^\beta = u_0^\beta + \alpha(u_1^\beta - u_0^\beta)$ . Если  $t_n \leq t \leq t_{n+1}$ ,  $x_i \leq x \leq x_{i+1}$ , то положим  $\psi_h(t, x) = u_\alpha^\beta$  при  $\alpha = (t - t_n)/\tau_n$ ,  $\beta = (x - x_i)/h$ .

Очевидно, что  $\min(u_0^0, u_0^1, u_1^0, u_1^1) \leq u_\alpha^\beta \leq \max(u_0^0, u_0^1, u_1^0, u_1^1)$ . Так как  $c_n^i$  ограничены, то  $\psi_h$  равномерно ограничены в  $\mathbb{N}$ . Кроме того, все  $\psi_h$  равностепенно непрерывны в  $\mathbb{N}$ . В самом деле, выберем точки  $(t, x), (t + dt, x + dx) \in \mathbb{N}$ . Пусть  $t_n \leq t \leq t_{n+1}, x_i \leq x \leq x_{i+1}$ ; можно считать, что  $t_n \leq t + dt \leq t_{n+1}, x_i \leq x + dx \leq x_{i+1}$ . Тогда  $dt = \tau_n d\alpha, dx = h d\beta$ . Рассмотрим разность

$$\Delta\psi_h = \psi_h(t + dt, x + dx) - \psi_h(t, x) = u_{\alpha+d\alpha}^{\beta+d\beta} - u_\alpha^\beta = \partial_\alpha u_\alpha^\beta d\alpha + \partial_\beta u_\alpha^\beta d\beta.$$

Но  $0 \leq \partial_\alpha u_\alpha^\beta \leq Z\tau_n, 0 \leq \partial_\beta u_\alpha^\beta \leq Zh$ . Поэтому  $|\Delta\psi_h| \leq Z(|dt| + |dx|)$ . Это и означает равностепенную непрерывность.

Возьмем выбранную выше последовательность  $h_j \rightarrow 0$  и рассмотрим соответствующую последовательность  $\psi_{h_j}(t, x)$ . Она равномерно ограничена и равностепенно непрерывна. По теореме Арцела, из нее можно выделить равномерно сходящуюся подпоследовательность. Пусть ее предел  $c(t, x)$ . Соответствующие индексы опять обозначим  $j$ . Ясно, что  $1 \leq c(t, x) \leq \sqrt{P/b} + \tilde{c}$ , не убывает по  $x$  и  $t$  и удовлетворяет (5). Назовем пару  $(\rho(t), c(t, x))$  обобщенным решением задачи (1)–(5).

## 9. ОБОБЩЕННОЕ РЕШЕНИЕ

Покажем, что  $c(t, x)$  действительно является обобщенным (слабым) решением уравнения диффузии (1) в множестве  $Y_\rho(T)$ . Пусть  $\varphi(t, x)$  – финитные функции в  $Y_\rho(T)$  ( $\varphi(t, x) \in C^\infty$  и носитель компактен в  $Y_\rho(T)$ ).

Возьмем одну  $\varphi(t, x)$  (пусть для определенности ее носитель – прямоугольник в  $Y_\rho(T)$ ) и выберем  $h$  таким малым, чтобы граничные узлы  $\bar{D}_N$  (те, которые входят в  $\bar{D}_N$ , но не входят в  $D_N$ ) не входили в носитель  $\varphi$ . Пусть в него входят узлы  $(n, i), n_- + 1 \leq n \leq n_+ - 1, i_- + 1 \leq i \leq i_+ - 1$ .

Введем обозначение  $\varphi_n^i = \varphi(t_n, x_i)$ . Рассмотрим сумму

$$\sum_{n=n_-}^{n_+ + 1} \sum_{i=i_- - 1}^{i_+ + 1} c_{n+1}^i (\partial_\tau \varphi_{n+1}^i + \partial_{hh} \varphi_{n+1}^i) h \tau_n. \quad (23)$$

Это интегральная сумма. С другой стороны, выполним операцию, аналогичную интегрированию по частям. При этом (23) равна

$$\sum_{n=n_-}^{n_+} \sum_{i=i_-}^{i_+} \varphi_n^i (-\partial_\tau c_{n+1}^i + \partial_{hh} c_{n+1}^i) h \tau_n = 0.$$

Равенство нулю — в силу (7), при всех  $h$ . Это означает, что

$$\int_{\bar{Y}_\rho(T)} c(t, x) \left( \partial_t \varphi + \partial_{xx} \varphi \right) dt dx = 0$$

для любой финитной в  $\bar{Y}_\rho(T)$  функции  $\varphi(t, x)$  (ведь любой компакт можно погрузить в объединение прямоугольников), то есть  $c(t, x)$  является обобщенным решением уравнения диффузии.

Докажем, что имеются непрерывные (по  $t$ ) производные  $\partial_x c(t, L)$  и  $\partial_x c(t, \rho(t))$ , причем граничные условия (2), (3) выполнены (в обычном смысле).

В самом деле, подставим  $\psi_h(t, x)$  в (3). Если  $K_{n+1} = K_n$ , то

$$\begin{aligned} \psi_h(t, \rho_h(t)) &= u_\alpha^{1-\alpha} = \\ &= c_n^{k_n} - 2\alpha c_n^{k_n} + \alpha c_{n+1}^{k_n} + \alpha c_n^{k_n-1} - \alpha^2 (c_{n+1}^{k_n} - c_n^{k_n} - c_{n+1}^{k_n-1} + c_n^{k_n-1}), \\ \partial_x \psi_h(t, \rho_h(t)) &= \partial_\beta \partial u_\alpha^\beta / h = \partial_h c_n^{k_n} + \alpha (\partial_h c_{n+1}^{k_n} - \partial_h c_n^{k_n}). \end{aligned}$$

Так как  $\partial_h c_n^{k_n} = \partial_h c_n^{k_n+1} = c_n^{k_n} - 1$ , то, подставляя в (3), имеем

$$\begin{aligned} -2c_n^{k_n} + c_{n+1}^{k_n} + c_n^{k_n-1} - \alpha (c_{n+1}^{k_n} - c_n^{k_n} - c_{n+1}^{k_n-1} + c_n^{k_n-1}) &= \\ &= \partial_h c_{n+1}^{k_n} - \partial_h c_n^{k_n} + \Psi_n / \alpha. \end{aligned}$$

Здесь  $\Psi_n$  — невязка. Применяем (9) еще раз ( $\alpha(1 - \alpha) < 1$ ):

$$|\Psi_n| \leq \left| c_n^{k_n} - c_{n+1}^{k_n} - c_n^{k_n-1} + c_{n+1}^{k_n-1} \right| \leq \left| \partial_h c_n^{k_n} - \partial_h c_{n+1}^{k_n} \right| h \leq Z h = O(h).$$

Если  $K_{n+1} = K_n + 1$ , то

$$\begin{aligned} \psi_h(t, \rho_h(t)) &= u_\alpha^0 = c_n^{k_n} + \alpha (c_{n+1}^{k_n} - c_n^{k_n}), \\ \partial_x \psi_h(t, \rho_h(t)) &= \partial_\beta \partial u_\alpha^\beta / h = \partial_h c_n^{k_n} + \alpha (\partial_h c_{n+1}^{k_n} - \partial_h c_n^{k_n}) \end{aligned}$$

и (3) выполняется точно. В любом случае

$$\partial_x \psi_h(t, \rho_h(t)) = \psi_h(t, \rho_h(t)) - 1 + O(h).$$

Правая часть равномерно сходится при  $h \rightarrow 0$ , ( $\rho_h$  тоже равномерно сходится), поэтому существует равномерный предел левой части (непрерывная производная) и условие (3) выполнено для построенного решения  $c(t, x)$ ,  $\rho(t)$ .

Теперь подставим  $\psi_h(t, x)$  в (2). При этом

$$\begin{aligned} \psi_h(t, L) &= u_\alpha^1 = c_n^I + \alpha (c_{n+1}^I - c_n^I), \\ \partial_x \psi_h(t, L) &= \partial_\beta \partial u_\alpha^\beta / h = \partial_h c_n^I + \alpha (\partial_h c_{n+1}^I - \partial_h c_n^I). \end{aligned}$$

Подставляя, учтем невязку  $\Psi_n$ :

$$\begin{aligned}\partial_h c_n^I + \alpha(\partial_h c_{n+1}^I - \partial_h c_n^I) + \Psi_n &= P - b(c_n^I + \alpha(c_{n+1}^I - c_n^I) - \tilde{c})^2 = \\ &= P - b(c_n^I - \tilde{c})^2 - b\alpha^2(c_{n+1}^I - c_n^I)^2 - 2b\alpha(c_{n+1}^I - c_n^I)(c_n^I - \tilde{c}).\end{aligned}$$

Применяем (8):

$$\Psi_n = b\alpha((c_{n+1}^I)^2 - 2c_{n+1}^I c_n^I + (c_n^I)^2 - \alpha(c_{n+1}^I - c_n^I)^2).$$

Отсюда получаем ( $\alpha(1 - \alpha) < 1$ )

$$\Psi_n \leq b(c_{n+1}^I - c_n^I)^2 = b(\partial_\tau c_{n+1}^I \tau_n)^2 = O(h).$$

Здесь мы применили утверждение 8 и то, что  $\tau_n \leq \bar{\tau} = O(\sqrt{h})$ . Итак, (6) выполняется на аппроксимациях с точностью  $O(h)$ :

$$\partial_x \psi_h(t, L) = P - b(\psi_h(t, L) - \tilde{c})^2 + O(h).$$

Правая часть равномерно сходится при  $h \rightarrow 0$ , поэтому существует равномерный предел левой части (непрерывная производная) и условие (2) выполнено для построенного решения  $c(t, x)$ .

Запишем условие Стефана (4) в интегральной форме и подставим туда аппроксимации  $\psi_h(t, x)$  и  $\rho_h(t)$ :

$$\rho_h(t) = \rho_0 - t + \int_0^t \psi_h^{-1}(\xi, \rho_h(\xi)) d\xi + \Psi_h. \quad (24)$$

Здесь  $\Psi_h$  — невязка. Мы хотим доказать, что  $\Psi_h \rightarrow 0$  при  $h \rightarrow 0$ .

Рассмотрим сначала  $\rho_h(t)$ . Пусть  $t_n \leq t \leq t_{n+1}$ . По построению

$$\rho_h(t_n) = x_{k_n} = x_K + \sum_{m=1}^n x_m - x_{m-1} = \rho_0 + \sum_{m=1}^n \partial_\tau \rho_m \tau_{m-1}.$$

Пусть  $\mu_n$  — множество тех  $m = 1, \dots, n$ , для которых  $c_m^{k_m}$  не мало, то есть  $c_m^{k_m} - 1 \geq B\sqrt{h}$ . Через  $\bar{\mu}_n$  обозначим дополнение, то есть

те  $m = 1, \dots, n$ , которые не вошли в  $\mu_n$ . Применим (12):

$$\begin{aligned} \rho_h(t_n) - \rho_0 &= \sum_{m \in \mu_n} \partial_\tau \rho_m \tau_{m-1} = \sum_{m \in \mu_n} \left( (c_m^{k_m})^{-1} - 1 \right) \tau_{m-1} = \\ &= \sum_{m=1}^n \left( (c_m^{k_m})^{-1} - 1 \right) \tau_{m-1} - \sum_{m \in \bar{\mu}_n} \left( (c_m^{k_m})^{-1} - 1 \right) \tau_{m-1} = \\ &= \sum_{m=1}^n \left( \tau_{m-1} (c_m^{k_m})^{-1} \right) - t_n + O(\sqrt{h}). \end{aligned}$$

Теперь заметим, что  $\rho_h(t) = \rho_h(t_n) + O(\tau_n)$ .

Рассмотрим

$$\int_0^{t_n} \psi_h^{-1}(\xi, \rho_h(\xi)) d\xi = \sum_{m=1}^n \int_{t_{m-1}}^{t_m} \psi_h^{-1}(\xi, \rho_h(\xi)) d\xi.$$

Подынтегральные функции непрерывны и ограничены, поэтому

$$\int_{t_{m-1}}^{t_m} \psi_h^{-1}(\xi, \rho_h(\xi)) d\xi = (c_m^{k_m})^{-1} \tau_{m-1} + o(\tau_{m-1}).$$

Перепишем последнюю  $o(\tau_{m-1}) = \tau_{m-1} \omega_{m-1}$ ,  $\omega_m \rightarrow 0$  при  $h \rightarrow 0$ .

Тогда рассматриваемый интеграл примет вид

$$\int_0^{t_n} \psi_h^{-1}(\xi, \rho_h(\xi)) d\xi = \sum_{m=1}^n (c_m^{k_m})^{-1} \tau_{m-1} + \sum_{m=1}^n \omega_{m-1} \tau_{m-1}$$

Последняя сумма стремится к нулю при  $h \rightarrow 0$ . Подставим найденные выражения в (24), объединяя все малые величины в  $W$ . Получим  $\Psi_h = W \rightarrow 0$  при  $h \rightarrow 0$ . Переайдем к пределу в (24). Получим, что непрерывное решение  $\rho(t)$ ,  $c(t, x)$  удовлетворяют условию Стефана в интегральной форме. Поэтому левая его часть (функция  $\rho(t)$ ) не только непрерывна, а также обладает непрерывной производной. Дифференцируя по  $t$ , получаем, что решение  $\rho(t)$ ,  $c(t, x)$  удовлетворяет исходному условию Стефана (4).

Подведем итоги. Мы построили обобщенное непрерывное положительное ограниченное монотонное по каждой переменной решение  $c(t, x)$  уравнения диффузии (1) и непрерывно-дифференцируемую монотонно убывающую подвижную границу  $\rho(t)$ ; решение  $c(t, x)$  обладает производными, входящими в граничные условия (3) и (2), и удовлетворяет им; условие Стефана (4)

и начальные условия (5) выполняются. Поэтому название пары  $(c(t, x), \rho(t))$  обобщенным решением задачи (1)–(5) обосновано.

### Литература

1. Сборник тезисов Международного симпозиума по металло-водородным системам (MH2004, Краков, Польша, сентябрь 2004).
2. Сборник тезисов IX международной конференции «Водородное материаловедение и химия углеродных наноматериалов» (ICHMS'05, Севастополь, Украина, сентябрь 2005, Киев, АНЕУ, 2005).
3. Заика Ю.В., Чернов И.А. Краевая задача с динамическими граничными условиями и движущейся границей (кинетика дегидрирования) // Математическое моделирование, Т. 16, № 4, 2004, с. 3-16.
4. Будак Б.М., Васильев Ф.П., Успенский А.Б. Разностный метод решения некоторых краевых задач типа Стефана // в кн.: Численные методы в газовой динамике, вып. 4 / под ред. Павлова Б.М.- М.: МГУ, 1965, с. 139–183.
5. Самарский А.А. Теория разностных схем.- М.: Наука, 1983.
6. Люстерник Л.А., Соболев В.И. Элементы функционального анализа.- М.: Наука, 1965.



**Раздел II. ИНФОРМАТИКА**

**Section II. COMPUTER  
SCIENCE**



УДК 519.87.872

**ДОВЕРИТЕЛЬНОЕ ОЦЕНИВАНИЕ ВЕРОЯТНОСТИ  
ПЕРЕПОЛНЕНИЯ БУФЕРА НА ОСНОВЕ УСКОРЕННОГО  
РЕГЕНЕРАТИВНОГО МОДЕЛИРОВАНИЯ СИСТЕМЫ  $M/M/1$**

А. В. БОРОДИНА, Е. В. МОРОЗОВ

В статье предложена модификация ускоренного метода имитационного моделирования Splitting для построения доверительного интервала вероятности переполнения буфера ( $\gamma$ ). Получена формула вероятности переполнения  $\gamma$  для системы  $M/M/1$  и проведено сравнение с оценкой вероятности  $\gamma$ , полученной методом Splitting.

Работа поддержана Российским фондом фундаментальных исследований, гранты 04-01-00671, 04-07-90115.

**1. ВВЕДЕНИЕ**

Широкий класс процессов, описывающих динамику коммуникативных сетей, имеет регенеративную структуру. Для эффективного интервального оценивания параметров таких систем используется регенеративный метод имитационного моделирования [1–3].

Поскольку оцениваемые вероятности, как правило, очень малы, то стандартные методы имитационного моделирования (Монте-Карло) не эффективны. Существует ряд методов ускоренного моделирования [4, 5], позволяющих получить оценку с заданной точностью с меньшими вычислительными затратами (в частности, метод Splitting и RESTART [6]).

В данной статье мы предлагаем некоторую модификацию метода Splitting [7–9] без отсечения траекторий (т. е. моделирование траектории останавливается только после достижения нулевого уровня). С помощью метода Splitting строится оценка вероятности достижения события на цикле регенерации (например,

вероятность преполнения буфера заявок до момента опустения системы обслуживания). Доверительный интервал оцениваемой вероятности строится на основе метода регенеративного моделирования.

## 2. МЕТОД РАСЩЕПЛЕНИЯ SPLITTING

Рассмотрим случайный процесс  $X = \{X(t), t \geq 0\}$  с пространством состояний  $E$ . Пусть множество  $A$  является подмножеством множества  $E$ . Далее будем называть множество  $A$  *множеством редких событий*. Требуется оценить вероятность  $\gamma = P(X \in A)$  редкого события из множества  $A \subset E$

$$\gamma = \int I_{(X \in A)} dF(x) = EI_{(X \in A)}, \quad (1)$$

где  $I_{(X \in A)}$  функция-индикатор редкого события из множества  $A$ .

**2.1. Оцениваемые вероятности.** Целью дискретного моделирования систем массового обслуживания обычно является построение оценки некоторых стационарных или переходных характеристик. Примером таких характеристик являются стационарное время ожидание заявок в очереди, вероятность потери заявок при переполнении, вероятность переполнения буфера до того, как система станет пустой и т.д.

Обозначим  $T_A$  момент первого достижения процессом множества  $A$ . Пусть  $T_B$  момент достижение множества  $B \subset E$  (обычно  $B = \{0\}\right).$  Далее  $T_B$  будет иметь смысл момента сильной регенерации для процесса  $X$ . (Например, для процесса изменения длины очереди в системе массового обслуживания  $T_A$  - момент переполнения буфера заявок,  $T_B$  - момент опустения очереди.) Тогда событие  $\{T_A < T_B\}$  отражает попадание в редкое множество  $A$  на цикле регенерации процесса  $X$ . Для вычисления оценки вероятности  $\gamma = P(T_A < T_B)$ , разделим пространство состояний  $E$  процесса  $X$  на  $M + 1$  вложенных подмножеств  $C_i$ :

$$A = C_{M+1} \subset C_M \subset \cdots \subset C_1. \quad (2)$$

Подмножества  $C_i$  задаются так называемой функцией значимости  $f$ , заданной на  $E$

$$C_i = \{x \in E \mid f(x) \geq T_i\}, \quad i \in [1, M + 1], \quad (3)$$

где  $\{T_i\}$  известные значения функции  $f$ , называемые порогами или уровнями. Вероятность  $\gamma$  представляет произведение условных вероятностей  $p_1 = P(C_1), \dots, p_{i+1} = P(C_{i+1}|C_i)$ ,  $i \in [1, M]$ ,  $\gamma \equiv p_1 p_2 \dots p_{M+1}$ .

Пусть  $\hat{p}_1, \dots, \hat{p}_{M+1}$  оценки для соответствующих условных вероятностей  $p_1, \dots, p_{M+1}$ , где каждая оценка  $\hat{p}_i$  вычисляется стандартным методом Монте-Карло. Ниже приведем описание алгоритма метода Splitting.

**2.2. Описание метода.** Обозначим количество достижений уровня  $T_i$  во время моделирования  $N_i$  ( $N_0 = 1$ ),  $i \in [1, M + 1]$  ( $N_{M+1}$  - количество редких событий  $\{T_A < T_B\}$ ). Пусть  $R_i$  задает количество подтраекторий, генерируемых на уровне  $T_i$ ,  $R_{M+1} = 1$  (на последнем уровне подтраектории не генерируются). Пошаговая схема алгоритма Splitting следующая:

- (1)  $N_0 = 1$ ,  $N_i = 0$  для всех  $i \in [1, M + 1]$ .
- (2) Моделировать траекторию процесса  $X$  стандартным методом Монте-Карло.
- (3) Моделировать текущую траекторию (сгенерированную на уровне  $i$ ). Если произошло достижение уровня  $T_{i+1}$ , то переходим к шагу 4. Если достигли уровня  $T_0$ , то переходим к шагу 2. Если произошло событие, завершающее моделирование, то перейти на шаг 5 алгоритма.
- (4) Если достигли уровня  $T_{i+1}$ , то увеличим счетчик  $N_i$  на 1, расщепим траекторию на  $R_{i+1}$  подтраекторий-потомков. Каждая траектория моделируется отдельно, согласно шагу 3.
- (5) Вычислим оценки условных вероятностей достижения уровня  $i$  при условии старта с уровня  $i - 1$ ,

$$\hat{p}_i = \frac{N_i}{R_{i-1} N_{i-1}}, \quad (4)$$

где  $R_{i-1} N_{i-1}$  представляет общее количество траекторий, стартующих с уровня  $i - 1$ . Перейти на шаг 6.

- (6) Результирующая оценка вероятности редкого события  $\{T_A < T_B\}$  равна

$$\hat{\gamma} = \prod_{i=1}^{M+1} \hat{p}_i. \quad (5)$$

*Замечание.* Необходимо отметить, что предложенная в данной работе модификация метода Splitting работает по циклам регенерации моделируемого процесса. Такая модификация первоначального алгоритма (см. [7–9]) позволяет применить аппарат регенеративного моделирования (который опирается на ЦПТ) для построения доверительного интервала оцениваемой вероятности  $\gamma$ .

Приведенный выше алгоритм реализован на языке C++ в виде рекуррентной функции, параметрами которой являются номер уровня на котором происходит расщепление и текущее значение функции значимости  $f(x)$ . Возврат из рекурсии происходит при достижении множества  $B$ .

### 3. АНАЛИТИЧЕСКАЯ ОЦЕНКА $P(T_A < T_B)$ ДЛЯ СИСТЕМЫ $M/M/1$

Рассмотрим односерверную модель системы массового обслуживания  $M/M/1$  с бесконечным буфером и независимо одинаково распределенными интервалами между приходами заявок  $\{\tau_n\}$ , имеющими распределение

$$P(\tau_n \leq x) = 1 - e^{-\lambda x},$$

где  $\tau_n = t_{n+1} - t_n$  и  $t_n$  - момент прихода заявки с номером  $n$ ,  $n \geq 1$ . Пусть н.о.р. времена обслуживания  $\{S_n\}$  имеют распределение

$$P(S_n \leq x) = 1 - e^{-\mu x}.$$

Рассмотрим случайный процесс  $X = \{\nu(t), t \geq 0\}$  на пространстве состояний  $E = \{0, 1, \dots\}$ , где  $\nu(t)$  - количество заявок в системе в момент времени  $t$ . Получим для системы  $M/M/1$  аналитическую оценку вероятности  $\gamma = P(T_A < T_B)$ . Для удобства определим два события

$$\begin{aligned} H_+ &= \{ \text{приход заявки произошел раньше, чем уход}\}, \\ H_- &= \{ \text{уход заявки произошел раньше, чем приход}\}. \end{aligned}$$

Пусть  $p = P(H_+)$  - вероятность увеличения длины очереди на 1 (скакка вверх),  $q = P(H_-)$  - вероятность уменьшения длины очереди на 1 (скакка вниз),  $L$  - заданное значение длины очереди (как правило  $L$  имеет смысл порога длины очереди, после которого происходит потеря заявок).

Пусть  $T_A$  - момент времени, когда длина очереди стала равна  $L$ ,  $T_B$  - момент опустения очереди.

Обозначим вероятность того, что длина очереди достигнет значения  $L$  до того, как очередь опустеет, при условии, что в момент времени  $t$  длина очереди равна  $n$

$$P_n(t) = P(T_A < T_B \mid \nu(t) = n), \quad (6)$$

где  $1 \leq n < L$ . По свойству марковости рассматриваемого процесса  $X$  левая часть (6) не зависит от  $t$ , поэтому далее вместо  $P_n(t)$  будем использовать  $P_n$ .

Определим граничные условия

$$P_L = P(T_A < T_B \mid \nu(t) = L) = 1, \quad P_0 = P(T_A < T_B \mid \nu(t) = 0) = 0. \quad (7)$$

Получим рекуррентное соотношение для  $P_n$

$$\begin{aligned} P_n &= P(T_A < T_B \mid \nu(t) = n) = \frac{P(T_A < T_B, \nu(t) = n)}{P(\nu(t) = n)} = \\ &= \frac{P(T_A < T_B, H_+, \nu(t) = n) + P(T_A < T_B, H_-, \nu(t) = n)}{P(\nu(t) = n)} = \\ &= \frac{P(T_A < T_B \mid H_+, \nu(t) = n) \cdot P(H_+, \nu(t) = n)}{P(\nu(t) = n)} + \\ &= \frac{P(T_A < T_B \mid H_-, \nu(t) = n) \cdot P(H_-, \nu(t) = n)}{P(\nu(t) = n)} = \\ &= P(T_A < T_B \mid \nu(t + \Delta t) = n + 1) \cdot P(H_+) + \\ &\quad P(T_A < T_B \mid \nu(t + \Delta t) = n - 1) \cdot P(H_-). \end{aligned} \quad (8)$$

Следовательно,  $P_n$  связаны следующим уравнением

$$P_n = p \cdot P_{n+1} + q \cdot P_{n-1}, \quad 1 \leq n < L. \quad (9)$$

Определим вероятности  $p$  и  $q$  для системы  $M/M/1$ , описанной выше. Рассмотрим две независимые последовательности н.о.в.  $\{\tau_n\}$  и  $\{S_n\}$ . Заявка с номером  $n$  ушла из системы (произошло событие  $H_-$ ), если  $S_n < \tau_n$  (т. е. время обслуживания заявки с номером  $n$  меньше времени до прихода следующей заявки с номером  $n + 1$ ). Следовательно вероятность  $q$  ухода заявки из системы  $M/M/1$  равняется

$$q = P(S_n < \tau_n) = \int_0^\infty \lambda e^{-\lambda t} (1 - e^{-\mu t}) dt = 1 - \frac{\lambda}{\lambda + \mu} = \frac{\mu}{\lambda + \mu}. \quad (10)$$

Соответственно формуле (10) вероятность прихода заявки в систему ( $P(H_+)$ ) равна  $p = 1 - q = \frac{\lambda}{\lambda + \mu}$ . Пусть  $\rho = \lambda/\mu$  - интенсивность траффика, тогда из (10) следует  $\rho = p/q$ .

Уравнение (9) также можно получить используя задачу о разорении игрока (см. [10]). Модифицируя формулы, приведенные в [10], получим, что разностное уравнение (9) с граничными условиями (7) имеет следующее решение. При  $\rho \neq 1$

$$P_n = \frac{\rho^L - \rho^{L-n}}{\rho^L - 1}. \quad (11)$$

Покажем, что (11) является решением для (9). Подставим формулу (11) в (9) и рассмотрим числитель дроби в правой части уравнения

$$\begin{aligned} p \cdot \rho^L - p \cdot \rho^{L-n-1} + q \cdot \rho^L - q \cdot \rho^{L-n+1} = \\ \rho^L \cdot (p + q) - \rho^{L-n} (\rho \cdot \rho^{-1} + q \cdot \rho) = \\ \rho^L - \rho^{L-n} = P_n (\rho^L - 1). \end{aligned} \quad (12)$$

В силу того, что  $\rho \neq 1$  выполняется  $\rho^L \neq 1$ , следовательно, формула (11) есть решение уравнения (9).

Рассмотрим  $\rho = 1$ . В этом случае формула (11) не имеет смысла. Получим решение из (11) переходя к пределу и воспользовавшись правилом Лопитала

$$\lim_{\rho \rightarrow 1} \frac{\rho^L - \rho^{L-n}}{\rho^L - 1} = \lim_{\rho \rightarrow 1} \frac{L \cdot \rho^{L-1} - (L-n) \cdot \rho^{L-n-1}}{L \cdot \rho^{L-1}} = \frac{n}{L}. \quad (13)$$

Оцениваемая вероятность  $\gamma = P(T_A < T_B)$  является вероятностью достижения уровня  $L$  при единичном начальном значении длины очереди, т. е.

$$\gamma = P(T_A < T_B \mid \nu(t) = 1) = P_1 = \begin{cases} \frac{1}{L}, & \rho = 1 \\ \frac{\rho^L - \rho^{L-1}}{\rho^L - 1}, & \rho \neq 1. \end{cases} \quad (14)$$

где  $\rho = \lambda/\mu$  - интенсивность трафика.

На рисунке ?? приведены результаты моделирования вероятности  $\gamma = P(T_A < T_B)$  для системы  $M/M/1$  методом Splitting и результаты, полученные по формуле (14). Функцией значимости для метода Splitting (см. часть 1) является  $f(t) = \nu(t)$ . Оценка для  $\gamma$  методом Splitting при фиксированных значениях  $\rho$  и  $L$  была получена на основе выборки объема 500 случайных величин.

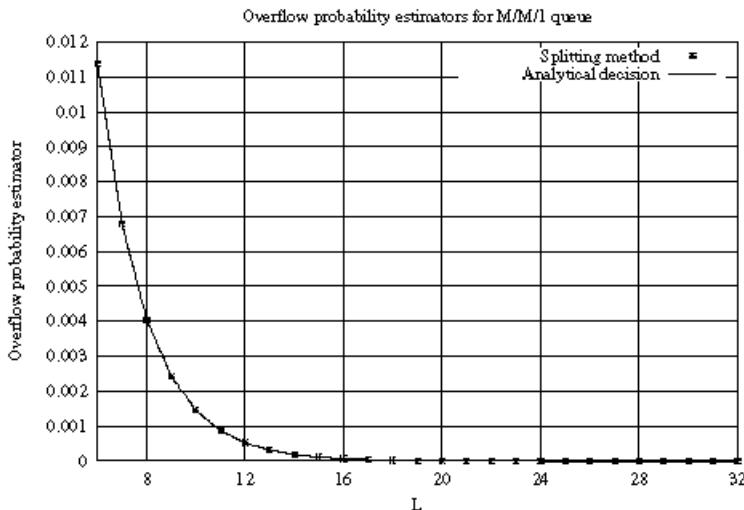


Рис. 1. Точечные оценки для вероятности  $\gamma$  в системе  $M/M/1$  при  $\rho = 0.6$

#### 4. РЕГЕНЕРАТИВНОЕ МОДЕЛИРОВАНИЕ

Классический метод регенеративного моделирования основан на использовании циклов сильной регенерации для эффективного интервального оценивания.

**4.1. Построение интервальных оценок.** Рассмотрим  $X = \{X(t), t \geq 0\}$  регенерирующий процесс (возможно с задержкой), заданный на пространстве состояний  $E$  с моментами регенерации  $T(-1) = 0 \leq T(0) < T(1) < \dots$ . Пусть  $\alpha_n = T(n) - T(n-1)$ ,  $n \geq 0$  - длина цикла регенерации. Пусть  $E\tau_1 < \infty$  (т. е. процесс  $X$  является положительно возвратным). Для заданной измеримой функции  $f : E \rightarrow \mathbb{R}$  определим последовательность случайных величин

$$Y_n(f) = \int_{T(n-1)}^{T(n)} f(X(s))ds, \quad n \geq 0. \quad (15)$$

**Условие 1.** Для регенеративного процесса  $X$  с  $E\alpha_1 < \infty$  существует константа  $\gamma$  такая, что выполняется  $E[Y_1 - \gamma\alpha_1] = 0$  и  $0 < E(Y_1 - \gamma\alpha_1)^2 < \infty$ .

Если процесс  $X$  является регенерирующим, то  $\{Z_n, n \geq 1\}$ , где  $Z_n = Y_n - \gamma\alpha_n$  - это последовательность независимых одинаково распределенных случайных величин. Условие 1 означает, что  $E|Z_1| < \infty$ ,  $E|Y_1| < \infty$  и  $\gamma = EY_1/E\alpha_1$ . Оценку для параметра  $\gamma$  можно получить моделированием на временном интервале  $(0, t)$  либо моделированием по  $n$  циклам регенерации. Будем рассматривать случай моделирования по циклам регенерации, тогда оценка для  $\gamma$  имеет следующий вид

$$\gamma_n = \frac{n^{-1} \sum_{k=1}^n Y_k}{n^{-1} \sum_{k=1}^n \alpha_k} = \frac{\bar{Y}_n}{\bar{\alpha}_n}. \quad (16)$$

Известно, что оценка  $\gamma_n$  для параметра  $\gamma$  будет строго состоятельной, если  $E|Y_1| < \infty$  и  $E\alpha_1 < \infty$ . Цель регенеративного метода состоит в построении асимптотически правильного доверительного интервала для  $\gamma$  (для построения используется ЦПТ).

**Теорема 1.** Пусть условие 1 выполняется, тогда имеет место слабая сходимость

$$n^{1/2}(\gamma_n - \gamma) \Rightarrow \sigma N(0, 1), \quad n \rightarrow \infty,$$

где  $\sigma^2 = EZ_1^2/(E\alpha_1)^2$ .

При построении доверительного интервала для  $\gamma$ , необходимо иметь слабо состоятельную оценку  $v_n$  дисперсии  $\sigma^2$ , тогда 100(1 –  $\delta$ )-процентный доверительный интервал строится по формулам

$$\left[ \gamma_n - \frac{z_\delta \sqrt{v_n}}{\sqrt{n}}, \gamma_n + \frac{z_\delta \sqrt{v_n}}{\sqrt{n}} \right], \quad (17)$$

где  $z_\delta$  - квантиль, определяемая вероятностью  $P(N(0, 1) \leq z_\delta) = 1 - \delta/2$ . Оценка  $v_n$  строится следующим образом

$$v_n = \frac{1/n \sum_{i=1}^n (Y_i - \gamma_n \alpha_i)^2}{\bar{\alpha}_n^2}, \quad n \geq 1. \quad (18)$$

**Теорема 2.** Если условие 1 выполняется, тогда оценка  $v_n$  является слабо состоятельной, т. е.  $v_n \Rightarrow \sigma^2$ ,  $n \rightarrow \infty$ .

Для строгой состоятельности оценки необходимы дополнительные условия  $EY_1^2 < \infty$  и  $E\alpha_1^2 < \infty$ . Однако для применимости регенеративного метода достаточно слабой сходимости оценок. Таким образом, условие 1 является достаточным условием применимости регенеративного метода. Более подробно см. [1, 2, 11].

### 5. ДОВЕРИТЕЛЬНОЕ ОЦЕНИВАНИЕ ВЕРОЯТНОСТИ $P(T_A < T_B)$

Рассмотрим регенерирующий процесс  $X = \{\nu(t), t \geq 0\}$  без задержки, где  $\nu(t)$  - это количество заявок в системе  $M/M/1$  в момент времени  $t$ . Для моделирования будем использовать  $n$  циклов регенерации процесса  $X$ . Пусть  $\alpha_i = T(i) - T(i-1)$ ,  $i \in [1, n]$  длины циклов регенерации с соответствующими моментами регенерации  $T(i)$ ,  $i \in [0, n]$ . Далее применим метод имитационного регенеративного моделирования (см. часть 3) для построения доверительного интервала для вероятности  $\gamma = P(T_A < T_B)$ .

*Замечание.* Поскольку процесс изменения длины очереди  $X$  для  $M/M/1$  марковский, то любое попадание процесса в фиксированное состояние  $\nu(t) = k$  является моментом регенерации ( $k$ -го типа). Модель  $M/M/1$  хорошо изучена, однако  $M/M/1$  и моменты регенерации 0-го типа (при  $k = 0$ ) используются в данной работе для иллюстрации метода и возможности сравнения результатов моделирования методом Splitting с аналитической формулой (14) (см. часть 2).

Пусть в формуле (15) заданная функция случайного процесса равна  $f = I_{\{T_A < T_B\}}^{(i)}$ , где индикатор  $I_{\{T_A < T_B\}}^{(i)} = 1$ , если на цикле регенерации с номером  $i$  произошло событие  $\{T_A < T_B\}$ . Следовательно

$$Y_n = \int_{T(i-1)}^{T_i} I_{\{T_A < T_B\}}^{(i)} ds = I_{\{T_A < T_B\}}^{(i)}(T(i) - T(i-1)) = I_{\{T_A < T_B\}}^{(i)} \alpha_i. \quad (19)$$

Тогда формулу (16) можно переписать в следующем виде

$$\gamma_n = \frac{\sum_{i=1}^n I_{\{T_A < T_B\}}^{(i)} \alpha_i}{\sum_{i=1}^n \alpha_i} = \frac{I_{\{T_A < T_B\}}^{(1)} \alpha_1 + \cdots + I_{\{T_A < T_B\}}^{(n)} \alpha_n}{\alpha_1 + \cdots + \alpha_n} \quad (20)$$

Числитель формулы (20) является суммой длин циклов регенерации, на которых произошло событие  $\{T_A < T_B\}$ . Знаменатель является суммарной длиной всех циклов регенерации.

С другой стороны рассмотрим оценку вероятности  $\gamma$ , построенную в методе Splitting. Перепишем формулу (5)

$$\hat{\gamma} = \frac{N_{M+1}}{R_0 R_1 \cdots R_M}. \quad (21)$$

Числитель формулы (21) представляет общее количество событий  $\{T_A < T_B\}$  (достижений уровня  $M + 1$ ). Знаменатель равен общему числу траекторий, полученных при моделировании методом Splitting. Каждая траектория является отдельным циклом регенерации процесса  $X$  (и могла быть аналогично построена методом классического моделирования Монте-Карло). Таким образом модификация метода Splitting (траектория обрывается только при достижении множества  $B$ ) позволяет сгенерировать  $R_0 \cdot R_1 \cdots R_M$  циклов регенерации. Следовательно в формуле (20) число циклов регенерации  $n = R_0 \cdot R_1 \cdots R_M$ .

Согласно свойству PASTA (Poisson Arrival See Time Average) (см. [12]) для систем, в которых моменты прихода заявок  $t_n$ ,  $n \geq 1$  распределены по закону Пуассона (т. е.  $M/\cdot/\cdot$ ), оценка по формуле (20) эквивалентна оценке, полученной методом Splitting (формула (21)) и выраженной в терминах событий (т. е.  $\hat{\gamma} = \gamma_n$ ), тогда

$$\sum_{i=1}^n \alpha_i = \frac{n \cdot \sum_{i=1}^n I_{\{T_A < T_B\}}^{(i)} \alpha_i}{N_{M+1}}. \quad (22)$$

Построим доверительный интервал для  $\gamma$  по формуле (17). Для этого получим оценку  $v_n$  для  $\sigma^2$  по формулам (18) и (22)

$$v_n = \frac{n \cdot \sum_{i=1}^n (I_{\{T_A < T_B\}}^{(i)} \cdot \alpha_i - \gamma_n \cdot \alpha_i)^2}{(\sum_{i=1}^n \alpha_i)^2} = \frac{N_{M+1}^2 \cdot \sum_{i=1}^n [\alpha_i^2 \cdot (I_{\{T_A < T_B\}}^{(i)} - \gamma_n)^2]}{n \cdot [\sum_{i=1}^n (I_{\{T_A < T_B\}}^{(i)} \cdot \alpha_i)]^2}. \quad (23)$$

Подставив формулу (23) в выражение (17), получим доверительный интервал для вероятности  $\gamma$ , оцениваемой методом Splitting.

*Замечание.* Для построения доверительного интервала с помощью регенеративного метода требуется внести изменения в описанный в части 1 алгоритм. В процессе моделирования необходимо вести подсчет длин циклов  $\alpha_i$ ,  $i \in [1, n]$ , и хранить значения  $I_{\{T_A < T_B\}}^{(i)}$ ,  $i \in [1, n]$  для каждого цикла регенерации.

## 6. ЗАКЛЮЧЕНИЕ

В данной статье предложена модификация метода ускоренного моделирования Splitting для построения доверительного интервала вероятности переполнения  $\gamma$ . Моделирование по циклам регенерации рассматриваемого процесса длины очереди позволяет строить интервальные оценки, используя регенеративный подход. Формулы, приведенные в части 4, позволяют строить доверительный интервал в процессе ускоренного моделирования методом Splitting.

Кроме того получены точечные оценки вероятности  $\gamma$  для системы массового обслуживания  $M/M/1$  и проведено сравнение с аналитическим решением (14). Рисунок ?? показывает, что оценка, полученная моделированием, практически совпадает с аналитической оценкой.

## Литература

1. P. W. Glynn, D. L. Iglehart. (1993). Conditions for the applicability of the regenerative method // Management Science 39: 1108-1111. <http://citeseer.ist.psu.edu/36432.html>
2. P. W. Glynn. Some topics in regenerative steady state simulation. 1993.
3. K. Sigman, R. Wolff. A review of regenerative processes. SIAM Review, Vol. 35, No. 2, pp.269-288. 1993.
4. P. E. Heegaard, A survey of Speedup simulation techniques. Workshop tutorial on Rare Event Simulation, Aachen, Germany. 1997.
5. P. Shahabuddin. Rare event simulation in stochastic models. Proceedings of the WSC 1995, IEEE Press., pp. 178-185.
6. A. Borodina, E. Morozov. Simulation of rare events with speed-up techniques: Splitting and RESTART. Proceedings of Finnish Data Processing Week at the Petrozavodsk State University (FDPW'2005), Vol. 7, 2006.
7. M. Garvels. The splitting method in rare event simulation. May 2000.
8. P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic. A look at multilevel splitting. In H. Niederreiter, editor, Monte Carlo and Quasi Monte Carlo Methods 1996, Lecture Notes in Statistics, volume 127, pages 99-108. Springer Verlag, 1996.
9. P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic. Splitting for rare event simulation: analysis of simple cases. Proceedings of the 1996 Winter Simulation Conference.
10. В. Феллер. Введение в теорию вероятностей и ее приложения. Том 1. М.:Мир, 1964.
11. P. W. Glynn, D. L. Iglehart. A joint central limit theorem for the sample mean and regenerative variance estimator. Annals of Operations Research 8, 1987, 41-55.
12. R. W. Wolff. Poisson Arrivals See Time Average. Opns. Res, 30, 223-231. 1982.

УДК 512.56

**ОБ ОДНОЙ МОДЕЛИ СЛУЧАЙНОГО ГРАФА  
"ИНТЕРНЕТ" ТИПА**

Н.С. КИСЕЛЕВА

Дано описание случайных графов, у которых степени вершин — независимые случайные величины, распределенные по закону  $P(D = d) = d^{-\tau} / \zeta(\tau)$ , где  $D$  — случайная величина, равная степени вершины графа,  $d = 1, 2, \dots$ ,  $\zeta(\tau) = \sum_{d=1}^{\infty} d^{-\tau}$  — дзета-функция Римана,  $\tau \in (2, 3)$  (конечное математическое ожидание и бесконечная дисперсия), а образование ребер происходит равновероятно. Предложено исследовать такие случайные графы для описания топологии Интернет. Получены предельные распределения для крайних членов вариационного ряда, составленного из неубывающей последовательности степеней вершин графа, также доказана локальная предельная теорема для суммы степеней вершин при  $N \rightarrow \infty$ , где  $N$  — число вершин в графе.

Работа выполнена при поддержке гранта РФФИ 05-01-00007а.

Случайные графы в качестве моделей для изучения основных характеристик топологии Интернет интенсивно используются уже несколько лет. Одной из таких моделей посвящена статья [1]. Пусть граф содержит  $N$  вершин, имеющих номера  $1, 2, \dots, N$ . Обозначим  $D_1, \dots, D_N$  последовательность независимых одинаково распределенных случайных величин, принимающих значения  $1, 2, \dots$ . Будем считать, что случайная величина  $D_i$  равна числу ребер, инцидентных  $i$ -ой вершине.

В статье [1] используется следующий закон распределения степеней вершин:

$$\mathbf{P}\{D_i \geq d\} = d^{-\tau+1},$$

где  $i = 1, \dots, N, d = 1, 2, \dots, \tau \in (2, 3)$ . Отсюда

$$\mathbf{P}\{D_i = d\} = d^{-\tau+1} - (d+1)^{-\tau+1}.$$

На первом этапе конструирования графа рассматриваются вершины с выходящими из них ребрами и не учитываются вершины на другом конце каждого ребра. Для такой модели вводится понятие полуребра. Полуребро — это ребро, выходящее из вершины с незакрепленной вершиной на другом конце. На следующем этапе данные полуребра равновероятно соединяются между собой, образуя ребра графа. Если сумма степеней вершин нечетна, то добавляется новая вершина с номером 0 и единичной степенью  $D_0 = 1$ .

Новая модель, представленная в данной работе, является случайным графом с аналогичным построением, но с иным законом распределения степеней:

$$\mathbf{P}\{D_i = d\} = \frac{1}{\zeta(\tau)d^\tau}, \quad (1)$$

где  $d = 1, 2, \dots, \tau \in (2, 3), \zeta(\tau) = \sum_{d=1}^{\infty} d^{-\tau}$  — дзета-функция Римана.

В статье показано, что новая модель сохраняет основные свойства известной модели, указанные в статьях [1] и [3]. При этом выдвигается предположение, что предложенная модель может оказаться более удобной для решения некоторых задач в силу более простого выражения  $\mathbf{P}\{D_i = d\}$ .

Из (1) следует, что при  $d \rightarrow \infty$

$$\mathbf{P}\{D_i \geq d\} = \frac{1}{\zeta(\tau)(\tau-1)} d^{-\tau+1} (1 + o(1)). \quad (2)$$

Распределение (1) не имеет конечной дисперсии и имеет конечное математическое ожидание. Обозначим  $D$  степень произвольной вершины графа, тогда

$$\mathbf{E} D = \sum_{d=1}^{\infty} d \cdot \frac{1}{\zeta(\tau)d^\tau} = \frac{\zeta(\tau-1)}{\zeta(\tau)}.$$

Обозначим  $F(x)$  функцию распределения случайной величины  $D$ . Тогда

$$F(x) = \mathbf{P}\{D < x\} = \begin{cases} 0 & \text{при } x \leq 1, \\ 1 - \frac{1}{\zeta(\tau)(\tau-1)} \lceil x \rceil^{-\tau+1} (1 + o(1)) & \text{иначе,} \end{cases}$$

где  $\lceil x \rceil$  — наименьшее целое  $n \geq x$ .

Пусть  $G(d_1, \dots, d_N)$  — множество графов с  $N$  вершинами, у которых первая вершина имеет степень равную  $d_1$ , вторая  $d_2$  и т.д. Обозначим общую сумму степеней вершин

$$l = \sum_{i=1}^N d_i.$$

Вероятность того, что случайный граф  $g$  с  $N$  вершинами принадлежит множеству  $G(d_1, \dots, d_N)$  равна:

$$\begin{aligned} \mathbf{P}\{g \in G(d_1, \dots, d_N)\} &= \mathbf{P}\{D_1 = d_1, \dots, D_N = d_N\} = \\ &= \mathbf{P}\{D_1 = d_1\} \cdot \dots \cdot \mathbf{P}\{D_N = d_N\}. \end{aligned}$$

Поскольку все соединения полуребер между собой в пары равновероятны, естественно считать, что все графы в множестве  $G(d_1, \dots, d_N)$  равновероятны. Так как все полуребра занумерованы, то существует  $(l-1)(l-3) \cdot \dots \cdot 1$  способов соединить их между собой. Значит, условная вероятность выбора конкретного графа  $g$  из этого множества равна:

$$\mathbf{P}\{g | g \in G(d_1, \dots, d_N)\} = \frac{1}{(l-1)(l-3) \cdot \dots \cdot 1}.$$

Тогда вероятность появления конкретного графа  $g \in G(d_1, \dots, d_N)$  равна:

$$\mathbf{P}\{g\} = \frac{\mathbf{P}\{D_1 = d_1\} \cdot \dots \cdot \mathbf{P}\{D_N = d_N\}}{(l-1)(l-3) \cdot \dots \cdot 1}.$$

В рассматриваемой модели нет ограничений на соединения полуребер в пары, поэтому нет ограничения на существование кратных ребер и петель.

В статье [1] доказана лемма:

**Лемма 1.** Пусть  $U$  и  $V$  два непересекающихся множества вершин. Обозначим

$$F = \sum_{i \in U} D_i, \quad G = \sum_{j \in V} D_j.$$

Если  $FG/N \rightarrow \infty$  с вероятностью, стремящейся к 1, при  $N \rightarrow \infty$ , то с вероятностью, стремящейся к 1, существует вершина из множества  $U$ , которая связана ребром с какой-нибудь вершиной из  $V$ .

Из этой леммы следует, что вершины, степень которых не меньше, чем  $N^{1/2+\varepsilon}$ , где  $\varepsilon > 0$ , соединены между собой хотя бы одним своим ребром с вероятностью, стремящейся к 1, следовательно, они образуют полный граф. Покажем ниже, что если выбрать  $\varepsilon$  так, чтобы выполнялось неравенство  $\varepsilon < (3-\tau)/2(\tau-1)$ , то такие вершины существуют.

Обозначим  $\Omega$  – множество вершин графа, у которых степени вершин удовлетворяют условию:  $D_i \geq N^{1/2+\varepsilon}$ , где  $i = 1, \dots, N$ . Тогда мощность множества  $\Omega$  определяется:

$$|\Omega| = \sum_{i=1}^N \mathbf{I}\{D_i \geq N^{1/2+\varepsilon}\},$$

где  $\mathbf{I}(A)$  – индикатор события  $A$ .

Применив теорему Муавра–Лапласа при  $N \rightarrow \infty$ , получим следующий результат:

$$\mathbf{P}\{|\Omega| = k\} = \frac{1 + o(1)}{\sqrt{2\pi N p(1-p)}} \cdot e^{-\frac{(k-Np)^2}{2Np(1-p)}},$$

где  $p = \mathbf{P}\{D \geq N^{1/2+\varepsilon}\} = (\zeta(\tau)(\tau-1))^{-1} \cdot (N^{1/2+\varepsilon})^{-\tau+1}(1+o(1))$ .

Нетрудно видеть, что  $Np \rightarrow \infty$ , следовательно, число вершин, удовлетворяющих заданному условию, стремится к бесконечности.

Далее построим вариационный ряд

$$D_{(1)} \leq D_{(2)} \leq \dots \leq D_{(N)},$$

расположив  $D_1, D_2, \dots, D_N$  в неубывающем порядке. Найдем предельные распределения для крайних членов этого ряда.

**Теорема 1.** Пусть  $N \rightarrow \infty$  и  $k$  — фиксированное натуральное число, тогда

$$\mathbf{P}\{D_{(k)} = 1\} \rightarrow 1.$$

ДОКАЗАТЕЛЬСТВО. Пусть  $F_k(x), k = 1, 2, \dots, N$ , обозначает функцию распределения  $k$ -го члена вариационного ряда (3), тогда

$$F_k(x) = \mathbf{P}\{D_{(k)} < x\} = 1 - \mathbf{P}\{D_{(k)} \geq x\}.$$

Ясно, что

$$\begin{aligned} \mathbf{P}\{D_{(k)} \geq x\} &= \\ \mathbf{P}\{D_1 \geq x, \dots, D_{(N)} \geq x\} &+ C_N^1 \cdot \mathbf{P}\{D_1 < x, D_2 \geq x, \dots, D_{(N)} \geq x\} + \\ &+ C_N^2 \cdot \mathbf{P}\{D_1 < x, D_2 < x, D_3 \geq x, \dots, D_{(N)} \geq x\} + \dots + \\ &+ C_N^{k-1} \cdot \mathbf{P}\{D_1 < x, \dots, D_{k-1} < x, D_k \geq x, \dots, D_{(N)} \geq x\} = \\ &= \mathbf{P}\{D_1 \geq x, \dots, D_{(N)} \geq x\} + \\ &+ \sum_{i=1}^{k-1} C_N^i \cdot \mathbf{P}\{D_1 < x, \dots, D_{(i)} < x, D_{i+1} \geq x, \dots, D_{(N)} \geq x\}. \end{aligned}$$

Из определения  $F_k(x)$ , следует, что  $\mathbf{P}\{D_{(k)} < x\} = 0$  при  $x \leq 1$ .  
Далее при  $x > 1$

$$\begin{aligned} 1 - F_k(x) &= \left( c[x]^{-\tau+1} (1 + o(1)) \right)^N + \\ &+ \sum_{i=1}^{k-1} \frac{N!}{(N-i)! i!} \cdot \left( 1 - c[x]^{-\tau+1} (1 + o(1)) \right)^i \left( c[x]^{-\tau+1} (1 + o(1)) \right)^{N-i}, \end{aligned}$$

где  $c = (\zeta(\tau)(\tau-1))^{-1}$ . Исходя из того, что  $x > 1$  и  $c < 1$ , получим, что  $1 - F_k(x) \rightarrow 0$  при  $N \rightarrow \infty$ . Отсюда следует, что  $\mathbf{P}\{D_{(k)} \geq x\} \rightarrow 0$  при  $x > 1$  и  $N \rightarrow \infty$ . Утверждение теоремы доказано.

**Теорема 2.** Пусть  $N \rightarrow \infty$  и  $k$  — фиксированное неотрицательное целое число, тогда

$$\mathbf{P}\{(\tau-1) \ln D_{(N-k)} - \ln N < x\} = \frac{\Gamma(k+1, ce^{-x})}{k!} (1 + o(1)), \quad (3)$$

где  $c = (\zeta(\tau)(\tau-1))^{-1}$  и  $\Gamma(m, x) = \int_x^\infty t^{m-1} e^{-t} dt$  — дополнительная гамма-функция.

ДОКАЗАТЕЛЬСТВО.

$$\begin{aligned}
 \mathbf{P}\{(\tau - 1) \ln D_{(N-k)} - \ln N < x\} &= \mathbf{P}\{D_{(N-k)} < (Ne^x)^{\frac{1}{\tau-1}}\} = \\
 &= \mathbf{P}\{D_1 < (Ne^x)^{\frac{1}{\tau-1}}, \dots, D_N < (Ne^x)^{\frac{1}{\tau-1}}\} + \\
 &+ \sum_{i=1}^k C_N^i \cdot \mathbf{P}\{D_1 < (Ne^x)^{\frac{1}{\tau-1}}, \dots, D_{N-i} < (Ne^x)^{\frac{1}{\tau-1}}, \\
 &\quad D_{N-i+1} \geq (Ne^x)^{\frac{1}{\tau-1}}, \dots, D_N \geq (Ne^x)^{\frac{1}{\tau-1}}\} = \\
 &= \sum_{i=0}^k \frac{N!}{(N-i)!i!} \left(1 - \frac{c}{Ne^x}\right)^{N-i} \cdot \frac{c}{(Ne^x)^i}
 \end{aligned}$$

Применяя формулу Стирлинга, получим равенство

$$\begin{aligned}
 \mathbf{P}\{D_{(N-k)} < (Ne^x)^{\frac{1}{\tau-1}}\} &= e^{-e^{-x}} \sum_{i=0}^k \frac{(ce^{-x})^i}{i!} + o(1) = \\
 &= \frac{\Gamma(k+1, ce^{-x})}{k!} (1 + o(1)),
 \end{aligned}$$

что и требовалось доказать.

**Следствие 1.** При  $k = 0$  предельное распределение для максимальной степени графа имеет вид:

$$\mathbf{P}\{(\tau - 1) \ln D_{(N)} - \ln N < x\} = e^{-e^{-x}} (1 + o(1)). \quad (4)$$

Для нахождения предельных распределений суммы степеней вершин можно воспользоваться теоремами 2.1.1, 2.2.2, 2.6.1 и 4.2.1 из [2]. Однаково распределенные независимые случайные величины  $D_1, D_2, \dots, D_N$  являются решетчатыми с максимальным шагом, равным единице, и функцией распределения (1). Образуем суммы:

$$S_N = \frac{D_1 + D_2 + \dots + D_N - A_N}{B_N},$$

где  $A_N$  и  $B_N$  - нормирующие постоянные.

Если функция распределения  $F_N(x)$  сумм  $S_N$  слабо сходится к некоторой функции распределения  $G(x)$ , то говорят, что

$F(x)$  притягивается к  $G(x)$ . Множество всех функций, притягивающихся к  $G(x)$ , называется областью притяжения закона  $G(x)$ .

Из теоремы 2.1.1 книги [2] следует, что распределение сумм  $S_N$  имеет в качестве предельного распределения устойчивый закон  $G(x)$ , и вид нормирующих постоянных при медленно меняющейся в смысле Карамата функции  $h(x) = 1$  будет следующим:

$$B_N = N^{1/(\tau-1)}, \quad A_N = N \cdot \mathbf{E} D,$$

где  $1 < \tau - 1 < 2$ .

Функция распределения  $F(x)$  принадлежит области притяжения устойчивого закона с показателем  $\tau - 1$  в силу теоремы 2.6.1 из [2]. Отсюда и из теорем 4.2.1 и 2.2.2 из [2], получим следующий результат:

**Теорема 3.** При  $N \rightarrow \infty$  справедливо:

$$\sup_k \left| N^{1/(\tau-1)} \mathbf{P}\{L_N = k\} - g\left(\frac{k \cdot \zeta(\tau) - N \cdot \zeta(\tau-1)}{N^{1/(\tau-1)} \zeta(\tau)}\right) \right| \rightarrow 0, \quad (5)$$

где  $L_N = \sum_{i=1}^N D_i$ ,  $g(x)$  – плотность устойчивого распределения  $G(x)$  с показателем  $\tau - 1$ , причем характеристическая функция  $\varphi(x)$  устойчивого распределения  $G(x)$  имеет вид:

$$\begin{aligned} \varphi(t) &= \\ &= \exp \left\{ i\gamma t + \frac{\Gamma(3-\tau)}{\zeta(\tau)(\tau-1)(\tau-2)} \cos \frac{\pi(\tau-1)}{2} |t|^{\tau-1} \left( 1 + i \frac{t}{|t|} \operatorname{tg} \frac{\pi(\tau-1)}{2} \right) \right\}. \end{aligned}$$

Таким образом, полученные результаты свидетельствуют о том, что рассмотренные свойства предложенной модели аналогичны соответствующим свойствам модели, изучавшейся в статьях [1] и [3].

## Литература

1. Reittu H., Norros I. On the power-low random graph model of massive data networks, Performance Evaluation, 55, 2004, 3-23.
2. Ибрагимов И.А., Линник Ю.В. Независимые и стационарно связанные величины. М., 1965.
3. Степанов М.М. О предельных распределениях степеней узлов в случайных графах "Интернет" типа // Труды ИПМИ КарНЦ РАН, Петрозаводск, 2005, вып.6, с.235-242.

УДК 519.144.1 + 519.16

## КОМПОЗИТНЫЙ ЭВОЛЮЦИОННЫЙ АЛГОРИТМ ДЛЯ ПОТОКОВОЙ ЗАДАЧИ ШТЕЙНЕРА.

В.Д. Кукин.

Обсуждается композитный эволюционный алгоритм для решения задачи Штейнера с потоками и зависящими от них весами. Приводятся основные результаты тестирования этого алгоритма.

Задача Штейнера с потоками и зависящими от них весами – одно из обобщений проблемы Штейнера [1]. Далее для краткости эта задача называется потоковой. Она относится к классу *NP*-трудных задач [2], для поиска решений которых обычно применяют приближенные методы, в частности, методы эволюционного моделирования. В статье рассматривается один из композитных эволюционных алгоритмов, разработанных для решения задачи.

### 1. ПОСТАНОВКА ЗАДАЧИ

Дадим комбинаторную постановку потоковой задачи Штейнера, используя терминологию теории графов. Введем следующие обозначения:

$G = (V, D)$  – ориентированный к корню плоский граф, где

$V$  – множество вершин и  $D$  – множество дуг орграфа.

$V = V_0 \cup V_s$ ,  $|V| = |V_0| + |V_s| = n + s$ .

$V_0$  – множество вершин орграфа, расположенных в фиксированных точках плоскости, называемых терминальными вершинами.

$i_0 \in V_0$  – корень дерева, или выделенная терминальная вершина, в которой находится сток.

$M_0 = \{m_i \in R^+ | i \in V_0 \setminus \{i_0\}\}$  – множество мощностей источников в терминальных вершинах.

$V_s$  — множество вершин орграфа, расположенных в дополнительных свободно размещаемых точках плоскости, называемых точками Штейнера (ТШ).

$X_0 = \{(x_i, y_i) | x_i, y_i \in R^+; i \in V_0\}$  — заданное множество координат терминальных вершин.

$X_s = \{(x_i, y_i) | x_i, y_i \in R^+; i \in V_s\}$  — множество координат ТШ.

$D = D_r$  — множество дуг орграфа, соединяющих его вершины в топологии  $\tau \in T$ .

$T$  — множество топологий, определяющих способы соединения вершин орграфа дугами.

Источники порождают потоки на дугах. В узлах (внутренних вершинах) выполняются условия сохранения потоков: суммируются потоки из заходящих дуг и образуют потоки на исходящих дугах. Если узел в терминальной вершине, то поток из нее тоже входит в эту сумму.

$f(q)$  — весовая функция;  $f(q_{(i,j)})$  — значение весовой функции, зависящее от величины  $q_{(i,j)}$  потока на дуге  $(i, j)$ , или вес этой дуги.

$l$  евклидова метрика;  $l_{(i,j)} = ((x_i - x_j)^2 + (y_i - y_j)^2)^{1/2}$  — длина дуги  $(i, j)$  орграфа.

$L_{(i,j)} = l_{(i,j)} \cdot f(q_{(i,j)})$  — взвешенная длина дуги  $(i, j)$ .

В комбинаторной постановке потоковая задача Штейнера — это следующий объект:

$$WSTP = < V_0, M_0, X_0, l, f(q), (\tau, X_s), S_L >,$$

где пара  $(\tau, X_s)$  — такое решение задачи, для которого при начальных условиях  $(V_0, M_0, X_0, l, f(q))$  и условиях сохранения потоков сумма взвешенных длин дуг орграфа минимальна:

$$S_L = \sum_{\tau \in T: (i,j) \in D_\tau} L_{(i,j)} \rightarrow \min,$$

Соответствующий паре  $(\tau, X_s)$  орграф  $G = (V, D_\tau)$  является деревом. После оптимизации координат ТШ оно становится деревом Штейнера (ДШ), т.е. относительно минимальным деревом для данной топологии  $\tau \in T$ . Ищется дерево, минимальное среди ДШ для всех топологий (МДШ).

Мы решаем потоковую задачу, рассматривая только деревья с полными топологиями [1], т.е. с максимальным числом ТШ ( $s = n - 2$ ). Число вершин всегда равно  $N = 2n - 2$ . Вырожденные

дуги (нулевой длины), которые могут получаться при оптимизации положения ТШ, мы считаем полноправными элементами деревьев. Это не нарушает общности подхода, зато позволило унифицировать представление топологии строкой фиксированной длины [3].

Потоковая задача рассматривается в приложении к проектированию транспортных сетей, для которых характерна положительная вещественная вогнутая весовая функция  $f(q)$ . Она обычно интерполируется следующей кусочно-линейной функцией:

$$f(q) = K(q) + Tr(q) \cdot q,$$

где:  $K(q), Tr(q)$  – ступенчатые (кусочно-постоянные) функции от потока. В транспортных сетях  $K(q)$  и  $Tr(q)$  принято называть капитальной и транспортной составляющими соответственно. В этом случае целевая функция задачи имеет вид:

$$S'_L = \sum_{\tau \in T: (i,j) \in D_\tau} [K(q_{(i,j)}) + Tr(q_{(i,j)}) \cdot q_{(i,j)}] \cdot l_{(i,j)} \rightarrow \min.$$

## 2. Композитный эволюционный алгоритм

На основе методологии эволюционного моделирования и нашего опыта решения потоковой задачи [3] были разработаны оригинальные генетические операторы и двухуровневая эволюционная модель. Они вместе с биологическими аналогиями приводятся в [4,5].

На верхнем уровне модели имитируется процесс случайного образования новых видов особей без их дивергенции. Особь представляется парой  $(\tau, X_s)$ , где  $\tau$  – топология (хромосома),  $X_s$  – координаты ТШ. На нижнем уровне моделируется эволюция популяции особей порожденного вида как смена ее поколений в дискретные моменты времени. Особь оценивается фитнес-функцией, соответствующей целевой функции потоковой задачи. За счет лучших найденных в популяциях разных видов особей формируется элитная группа решений.

На этой модели строились композитные алгоритмы, использующие различные приемы эволюционного моделирования. Вычис-

литературный эксперимент показал, что все они обеспечивают поиск хороших решений для широкого диапазона условий. Это достигается за счет применения оригинальных генетических операторов, учитывающих специфику задачи. Было отобрано несколько алгоритмов, имеющих лучшие временные характеристики, том числе рассматриваемый ниже алгоритм. Это алгоритм эволюционного поиска, в котором популяция состоит из одной особи, и он может интерпретироваться как стратегия эволюционного поиска.

Внешним циклом случайного видеообразования в алгоритме управляет параметр  $Z$ , который задает число актов образования новых видов. Первый шаг работы алгоритма – генерация родителя популяции. Первый родитель – это прародитель: автоматически формируемая особь предкового вида [4], а каждый следующий – случайный родитель, порождаемый из предыдущего.

Во внешний цикл вложен цикл эволюции популяции порожденного вида, которая моделируется как чередующиеся пары смен поколений – мутантов и гибридных особей. Этот цикл управляет критериальным признаком  $H$ . Перед каждой новой парой смены поколений признак  $H = 0$ . Если хотя бы в одном из двух последовательных поколений найден лучший потомок,  $H = 1$ .

Основные шаги алгоритма включают две последовательные операции: перестройку хромосомы (топологии  $\tau$ ) и локальную оптимизацию множества  $X_s$  координат ТШ на плоскости. Для перестройки используются: оператор случайных мутаций, оператор селективных мутаций и оператор кроссинговера [4]. Действиями этих операторов управляют специальные параметры – нормы, установленные в ходе вычислительного эксперимента.

### **Композитный эволюционный алгоритм.**

Шаг 1. Цикл случайного видеообразования:  $z = 1 \div Z$

Если  $z = 1$ , то генерация прародителя.

Иначе генерация случайного родителя вида  $z$ .

Перестройка хромосомы предыдущего родителя  
оператором случайных мутаций.

Оптимизация координат ТШ родителя.

Запоминание родителя.

Задание текущей особи.

Шаг 2. Эволюция популяции особей вида  $z$ .

Номер поколения популяции  $h = 1$ .

2.1. Критериальный признак  $H = 0$ .

Цикл генерации мутантов текущей особи:  $v = 2 \div (2n - 2)$

Порождение потомка  $v$  от текущей особи оператором селективных мутаций.

Оптимизация координат ТШ.

Если потомок лучше, текущая особь обновляется:

$H = 1$ .

Номер поколения популяции  $h = h + 1$ .

2.2. Цикл генерации гибридных особей:  $v = (n + 1) \div (2n - 2)$

Порождение потомка  $v$  от текущей особи оператором кроссинговера.

Оптимизация координат ТШ.

Если потомок лучше, текущая особь обновляется:

$H = 1$ .

Номер поколения популяции  $h = h + 1$ .

Если  $H = 1$ , переход к шагу 2.1.

Шаг 3. Элитная группа  $G$ .

Если  $G$  не создана, текущая особь включается в  $G$ ,

иначе, если текущая особь лучше худшей особи в  $G$ , заменяет ее.

Шаг 4. Если  $z < Z$ , то переход к шагу 1, иначе — к шагу 5.

Шаг 5. Выбор решения.

В цикле  $v = 2 \div (2n - 2)$  из текущей особи генерируется поколение мутантов с помощью оператора селективных мутаций. Геометрический смысл действия оператора — попытки перебросить дугу  $v$  на дуги дерева из ее ближайшей окрестности до первой удачной попытки или до исчерпания сформированного списка перебросок. Оператор применяется ко всем дугам, кроме корневой.

В цикле  $v = (n + 1) \div (2n - 2)$  порождается поколение гибридных особей с помощью оригинального оператора кроссинговера. Он реализован так, что в рекомбинации участвуют хромосомы

текущей особи и некоторой виртуальной особи. Для этого используется алгоритм исчерпывающего и неизбыточного перебора полных топологий [3].

Если при поиске найден лучший потомок (либо мутант, либо гибридная особь), текущая особь обновляется. Только она "выживает", и с нее продолжается поиск. Фактически, это алгоритм направленного эволюционного поиска.

В силу случайности видеообразования в алгоритме, не стоит проблема преждевременной сходимости. Он также просматривает большое число вариантов, что увеличивает вероятность нахождения хороших решений. Вместе с оригинальными операторами, учитывающими специфику потоковой задачи, это обеспечивает эффективность работы алгоритма, что подтверждается результатами его тестирования.

### 3. Вычислительный эксперимент

Вычислительный эксперимент состоял из тестирования разных реализованных на С++ алгоритмов, которое проводилось в основном на ПК с тактовой частотой 800 МГц. Искались решения как потоковой задачи Штейнера, так и задачи без потоков. Решались серии по 15 тестовых задач для каждой из размерностей: 50, 60, 70, 80, 90, 100, 250, 500 и 1000 терминальных вершин. Для задачи без потоков использовались тесты из Интернет-библиотеки [6], входные данные которых были получены с помощью генератора случайных чисел. Для потоковой задачи в них дополнительно задавались мощности источников в терминальных вершинах и весовая функция. Кроме того, решались потоковые задачи, для которых были известны решения из практики проектирования транспортных сетей. Для их входных данных характерна некоторая регулярность.

При тестировании рассмотренного алгоритма получены следующие результаты.

1. Алгоритм находит решения обеих задач за время, приемлемое для практического приложения к оптимизации параметров транспортных сетей. Он находит решения хорошего качества для широкого диапазона входных данных. Для задачи Штейнера без потоков найдены рекордные значения целевой функции, указанные в Интернет-библиотеке. Для потоковой задачи повторены

или улучшены решения, известные из практики проектирования.

2. Алгоритм способен находить субоптимальные решения. На рисунках ниже приведены схемы для задачи без потоков (рис.1) и соответствующей потоковой задачи (рис.2) размерностью 250 терминальных вершин с корнем в точке 1.

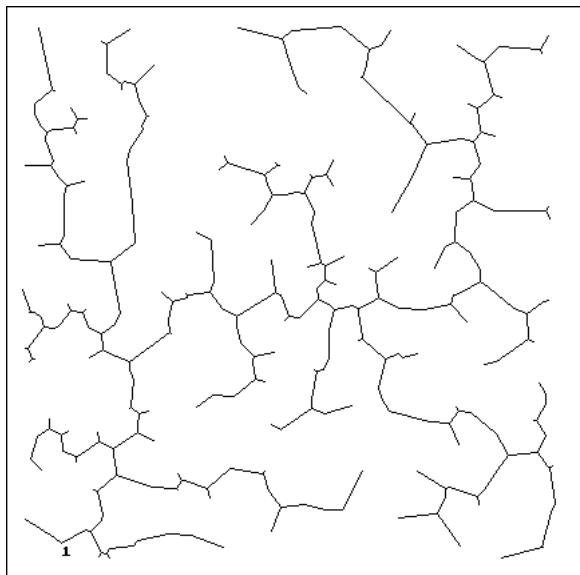
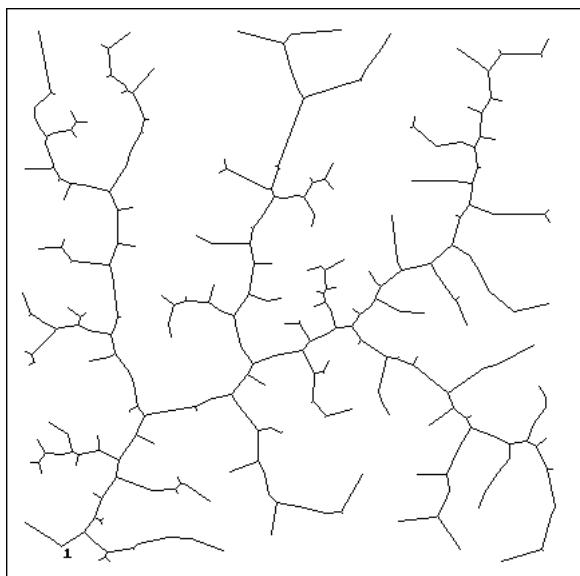


Рис. 1. Схема ДШ с 250 терминальными вершинами.  
Задача без потоков.

Для решений потоковой задачи характерны следующие особенности.

— В точках Штейнера углы между входящими дугами с близкими весами всегда меньше 120 градусов. Углы между входящими дугами с большой разницей весов близки к 90 градусам. (В задаче без потоков все углы равны 120 градусам.)

— Вероятность появления вырожденных дуг в потоковой задаче меньше, чем в задаче без потоков.



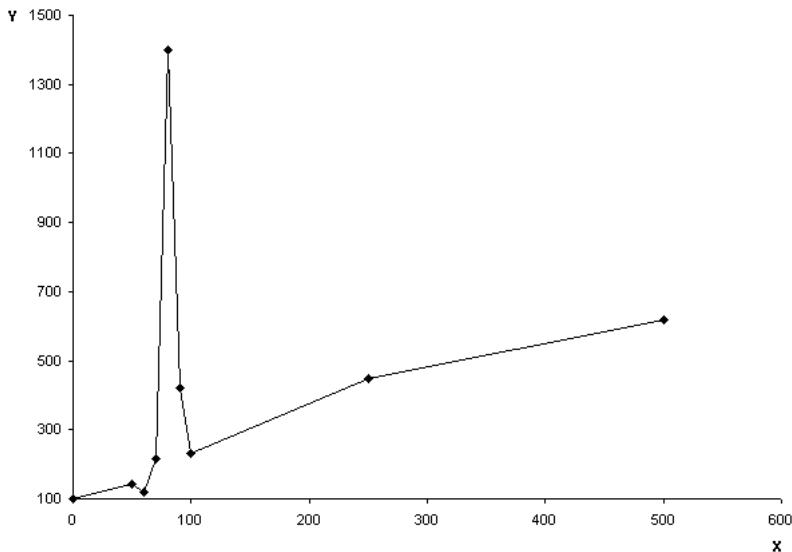
**Рис. 2.** Схема ДШ с 250 терминальными вершинами.  
Задача с потоками.

– ДШ с лучшим значением целевой функции в основном имеют большую сумму длин дуг, чем ДШ с хорошей целевой функцией – за счет увеличения длин дуг с меньшими весами.

– ДШ обладают относительной устойчивостью. Среди хороших деревьев, как правило, наблюдается малое число основных схем расположения дуг в прикорневой области, где дуги имеют больший вес. Различия начинаются ближе к периферии деревьев, где дуги имеют меньший вес. (Такая устойчивость совсем не свойственна задаче без потоков).

3. Алгоритм создавался для потоковой задачи и при поиске решений для нее более эффективен, чем для задачи без потоков. Для нахождения лучших решений в потоковой задаче всегда требовалось меньшее число актов видообразования, чем в задаче без потоков. Наличие весов порождает устойчивость решений в прикорневой области и в какой-то степени "сужает" пространство решений потоковой задачи. Это демонстрирует график на рис. 3. Он построен для контрольных тестовых задач разных размерностей как с потоками, так и без них. Показана относительная

разница числа актов видеообразования в них. Отметим, что в абсолютном выражении наблюдался одновременный резкий скачок числа видеообразований в задачах размерности 80. Обратная картина – для задач размерности 100. Это объясняется особенностями расположения терминальных вершин в этих примерах.



**Рис. 3.** Сравнение числа актов видеообразования.

$X$  – размерность задачи.

$Y$  – относительное число видеообразований в задаче без потоков.

Для потоковой задачи число видеообразований принято за 100% .

В обеих задачах наблюдается примерно одинаковый диапазон числа поколений популяции – от 4 до 24. Минимальное и максимальное число поколений наблюдается крайне редко – до 1% случаев. Наиболее вероятное значение также совпадает. В задачах размерностью 50-100 терминальных вершин чаще всего встречаются 8; в задачах размерностью 100-250 вершин – 10; в задачах размерностью 250-500 – 12 поколений. Однако максимальный процент наиболее вероятных значений числа поколений в потоковой задаче меньше, чем в задаче без потоков.

4. Для сравнения, на эволюционной модели был также построен алгоритм, в котором цикл эволюции популяции реализован по

схеме стандартного генетического алгоритма, но с нашими оригинальными операторами. Решались те же самые тестовые задачи — с потоками и без них. При поиске решений обеих задач этот алгоритм оказался менее эффективен, чем алгоритм направленного эволюционного поиска. Процесс затягивался и не гарантировал устойчивого нахождения решений. Это результат случайного действия генетических операторов, которое приводит к "хаотическому" блужданию по пространству решений. Так, в контрольной потоковой задаче на 80 терминальных вершин алгоритм направленного эволюционного поиска нашел хорошее решение за 9 актов видеообразования, а этот алгоритм — за 365. Поиск лучшего решения этим алгоритмом был прекращен, когда время превысило практически приемлемые пределы. Кроме того, примерно в 50% случаев эволюция популяции сводится к 1 поколению. Вследствие этого теряется качественное разнообразие популяции.

5. Тестировался также алгоритм, в котором число поколений эволюционирующей популяции фиксированное. Его производительность зависит от вложенных циклов. Сложность формирования поколения мутантов  $O(n^2)$ . В этом цикле  $2n - 4$  раз выполняется конечное число операций по перестройке хромосомы и оптимизируется положение  $n - 2$  пар координат ТШ. Аналогично, сложность формирования поколения гибридных особей  $O(n^2)$ . Временная сложность цикла эволюции популяции  $O(1)$ . При нахождении решения обеих задач максимальное число актов случайного видеообразования не превышало  $kn$ , где константа  $k \ll n$ . Значит, временная сложность этого цикла  $O(n)$ , а времененная сложность алгоритма с фиксированным числом поколений популяции  $O(n^3)$ . По-видимому, рассмотренный выше алгоритм имеет такую же временную сложность, ибо эмпирические данные свидетельствуют, что ожидаемое число поколений популяции с увеличением размерности задачи растёт очень слабо.

Таким образом, вычислительный эксперимент показал, что композитный эволюционный алгоритм имеет практическую ценность при реализации на ПК. В связи с этим, общий случай композитного алгоритма, в котором размер популяции равен числу процессоров многопроцессорной системы, перспективен для построения параллельного алгоритма.

#### 4. ЗАКЛЮЧЕНИЕ

Для решения потоковой задачи Штейнера использовалась методология эволюционного моделирования. На ее основе разработан ряд алгоритмов, в том числе алгоритм направленного эволюционного поиска, который и был рассмотрен выше.

Тестирование этого алгоритма позволило сделать следующие выводы.

— Алгоритм обеспечивает устойчивый поиск хороших и субоптимальных решений потоковой задачи и задачи Штейнера без потоков, но для последней работает медленнее.

— Алгоритм более эффективен для поиска решений задач как с потоками, так и без потоков, по сравнению с алгоритмом, в котором цикл эволюции популяции был реализован по схеме стандартного генетического алгоритма, но с использованием оригинальных операторов.

— Общий случай алгоритма, когда популяция состоит из конечного сравнительно небольшого числа особей, перспективен для построения параллельного алгоритма для многопроцессорной системы.

Таким образом, вычислительный эксперимент убедительно показал, что эволюционный подход к решению потоковой задачи Штейнера позволяет создавать алгоритмы, имеющие практическую ценность. Хотя они не гарантируют нахождения оптимального решения, они с большей вероятностью, чем методы последовательного поиска, способны находить субоптимальные решения.

#### Литература

1. Гилберт Э.Н., Поллак Г.О. Минимальные деревья Штейнера // Кибернетич. сборник, новая серия, вып. 8. М.: Мир, 1971, с. 19-50.
2. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982.
3. Кукин В.Д., Кузина В.И. Методика решения задачи Штейнера с потоками и весами // Труды ИПМИ КарНЦ РАН. Вып. 2. Петрозаводск, 2000, с. 143-150.
4. Кукин В.Д. Генетические алгоритмы и задача Штейнера с потоками и весами: подход к проблеме // Труды ИПМИ КарНЦ РАН. Вып. 3. Петрозаводск, 2002, с. 170-177.
5. Кукин В.Д. Эволюционная модель для решения потоковой задачи Штейнера // Труды ИПМИ КарНЦ РАН. Вып. 5. Петрозаводск, 2004, с. 200-211.
6. OR-LIBRARY: <http://mscmga.ms.ic.ac.uk/info.html>

УДК 681.142.2 + 519.2

## ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ТРЕМЯ СТЕКАМИ В СЛУЧАЕ ПАРАЛЛЕЛЬНОГО ВЫПОЛНЕНИЯ ОПЕРАЦИЙ

А. А. Лазутина

В работе предложены и исследуются математические модели, описывающие несколько методов представления трех стеков в памяти одного уровня в случае параллельного выполнения операций. В качестве критерия оптимальности рассмотрено среднее время до переполнения памяти.

Работа выполнена при поддержке гранта РФФИ (проект № 06-01-00303).

Понятие стека широко используется при разработке программного и аппаратного обеспечения [1]. При этом часто возникает проблема размещения нескольких стеков (процессов) в памяти одного уровня – оперативной или регистровой [2].

Существует несколько способов организации работы со стеками. Рассмотрим задачу об оптимальном способе размещения трех стеков в памяти одного уровня в случае параллельного выполнения операций.

Пусть в каждый момент времени могут происходить параллельные включения или исключения в три стека в памяти одного уровня.

В памяти объема  $m \geq 0$  условных единиц расположены три параллельных стека, вероятностные характеристики которых следующие:  $p_i$  – вероятность включения в стек с номером  $i$ ;  $q_i$  – вероятность исключения из стека с номером  $i$ ,  $p_{ij}$  – вероятность одновременного включения в два стека с номерами  $i$  и  $j$ ;  $q_{ij}$  – вероятность одновременного исключения из двух стеков с номерами  $i$  и  $j$ ;  $p_{123}$  – вероятность включения в три стека одновременно;  $q_{123}$  – вероятность исключения из трех стеков одновременно, где

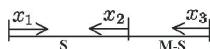
$p_i \geq 0, q_i \geq 0, i = 1, 2, 3$ . Исключение из пустого стека с вероятностью  $q_i$ ,  $q_{ij}$ ,  $q_{123}$  означает остановку на месте.

Предполагается, что  $p_1 + p_2 + p_3 + q_1 + q_2 + q_3 + p_{12} + p_{23} + p_{13} + q_{12} + q_{23} + q_{13} + p_{123} + q_{123} = 1$ .

Текущие длины стеков обозначим через  $x_1, x_2, x_3$ .

## 1. ПОСЛЕДОВАТЕЛЬНОЕ ПРЕДСТАВЛЕНИЕ

Пусть в памяти объема  $m$  единиц расположены три стека последовательно. Двум стекам, растущим навстречу друг другу, отведено  $s$  единиц памяти, а третьему стеку —  $m - s$  единиц (см. Рис. 1).



**Рис. 1.** Последовательное представление трех стеков в памяти одного уровня.

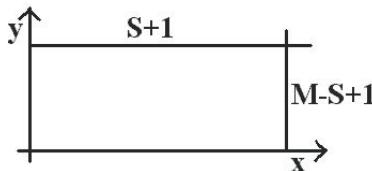
**1.1. Три стека, допускающих только включения.** Для начала рассмотрим более простой случай, когда возможны только включения в стеки.

Будем считать, что в каждый момент с разными вероятностями могут произойти следующие события: с вероятностью  $p_0$  все стеки остаются на месте, с вероятностью  $p_i$  произойдет включение только в стек с номером  $i$ , с вероятностью  $p_{12}$  произойдет операция параллельного включения в стеки 1 и 2, с вероятностью  $p_{13}$  произойдет операция параллельного включения в стеки 1 и 3, с вероятностью  $p_{23}$  произойдет операция параллельного включения в стеки 2 и 3, с вероятностью  $p_{123}$  произойдет операция параллельного включения во все три стека.

Тогда в зависимости от того какой стек будет расположен отдельно, мы будем иметь разные вероятности переходов. Пусть стек один и два объединены в пару, а третий стекложен отдельно. Введем обозначения:  $x = x_1 + x_2, y = x_3, r = p_0, p_1 = p_1 + p_2, p_2 = p_3, p_3 = p_{13} + p_{23}, p_4 = p_{12}, p_5 = p_{123}$ .

Тогда в качестве математической модели мы будем иметь двумерное случайное блуждание внутри прямоугольника [3],  $x \geq 0, y \geq 0, x < s + 1, y < m - s + 1$  с вероятностями переходов из точки  $(x, y)$  в точку  $(x, y)$  равной  $r$ , из точки  $(x, y)$  в точку  $(x + 1, y)$  равной  $p_1$ , в точку  $(x, y + 1)$  равной  $p_2$ , в точку  $(x + 1, y + 1)$  равной

$p_3$ , в точку  $(x + 2, y)$  равной  $p_4$ , в точку  $(x + 2, y + 1)$  равной  $p_5$  (см. Рис. 2).



**Рис. 2.** Область блуждания для последовательного представления трех стеков в памяти одного уровня в случае, когда происходят только включения.

Процесс выходит из начала координат и поглощается на прямых  $x = s + 1$ , и  $y = m - s + 1$ . Нашей задачей является нахождение такого значения  $s$  и определение того, какой стек расположить отдельно, чтобы среднее время блуждания до поглощения было максимальным. Таким образом здесь мы можем управлять только параметрами области, а не начальной точкой блуждания.

Обозначим через  $T(x, y)$  среднее время блуждания до поглощения, если начальной точкой блуждания была точка  $(x, y)$ . Тогда  $T(x, y) = r * T(x, y) + p_1 * T(x + 1, y) + p_2 * T(x, y + 1) + p_3 * T(x + 1, y + 1) + p_4 * T(x + 2, y) + p_5 * T(x + 2, y + 1) + 1$ .

Попав на границы области блуждания, мы уже не двинемся ни на один шаг и, следовательно,  $T(s + 1, y) = 0$ ,  $T(x, m - s + 1) = 0$ ,  $T(x, m - s + 2) = 0$ .

Нашей задачей является выбор стека, который будет расположен отдельно, и нахождение такого значения  $s$ , чтобы  $T(0, 0)$  было максимальным.

Был разработан алгоритм и написана программа на C++. Задавая вероятности и  $m$ , находится  $s$  и  $T(0, 0)$ .

Результаты для  $m = 100$  представлены в Таблицах 1-2.

**Таблица 1.**  $p_0 = p_1 = p_2 = p_3 = p_{12} = p_{13} = p_{23} = p$ ,  $p_{123} = 1 - 7 * p$ .

$p$	отдельно №1		отдельно №2		отдельно №3	
	$T(0, 0)$	$S$	$T(0, 0)$	$S$	$T(0, 0)$	$S$
0.05	41.542580	66	41.542580	66	41.542580	66
0.1	54.348526	66	54.348526	66	54.348526	66

Из таблицы 1 видно, что если все вероятности равны, то  $s$  и  $T(0, 0)$  не зависят от того, какой стек располагается отдельно.

**Таблица 2.**  $p_0 = p_1 = p_2 = p_{12} = p_{13} = p_{23} = p$ ,  $p_3 = 1 - 6 * p$ ,  
 $p_{123} = 0$ .

$p$	отдельно №1		отдельно №2		отдельно №3	
	$T(0, 0)$	$S$	$T(0, 0)$	$S$	$T(0, 0)$	$S$
0.05	87.295486	84	87.295486	84	85.757118	29
0.1	79.837135	73	79.837135	73	79.023659	50
0.15	73.716087	64	73.716087	64	73.850960	68

Вероятность включения во все три стека одновременно равна нулю,  $p_3$  выражается через остальные вероятности. Из Таблицы 2 видно, что при маленьком  $p$  большая вероятность включения в третий стек, поэтому под него отводится большая часть памяти; при большом  $p$  маленькая вероятность включения в третий стек, поэтому под него отводится меньшая часть памяти.

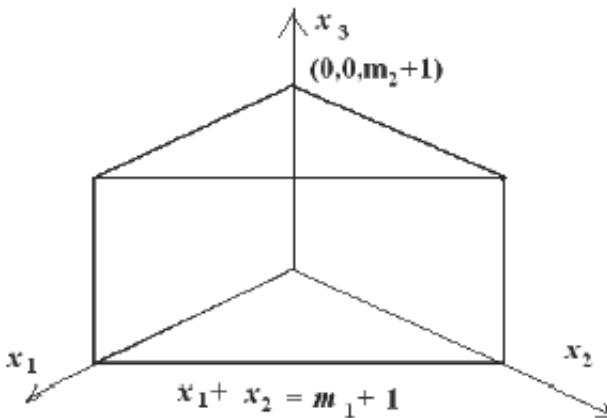
**1.2. Общий случай.** Пусть в памяти объема  $m \geq 0$  условных единиц расположены три параллельных стека последовательно (см. Рис. 1).

Обозначим  $m_1 = s$ ,  $m_2 = m - s$ .

Задача эффективного размещения трех стеков в памяти сводится к определению точки разбиения памяти на участки для пары стеков и для отдельно расположенного стека, т. е. к определению  $m_1$  и  $m_2$  так, чтобы среднее время работы до переполнения было бы максимальным. Дополнительно возникает вопрос, какой из стеков располагать отдельно.

В качестве математической модели задачи рассмотрим случайное блуждание точки, координаты которой определяются текущими длинами стеков  $(x_1, x_2, x_3)$ , по целочисленной решетке в пространстве внутри треугольной призмы с вершинами первого основания в точках  $(-1, -1, -1)$ ,  $(-1, m_1 + 1, -1)$ ,  $(m_1 + 1, -1, -1)$  и вершинами второго основания в точках  $(-1, -1, m_2 + 1)$ ,  $(-1, m_1 + 1, m_2 + 1)$ ,  $(m_1 + 1, -1, m_2 + 1)$  со следующими вероятностями сдвигов во внутренних точках:  $p_1, p_2, p_3$  — вероятности сдвига в положительном направлении вдоль осей  $x_1, x_2, x_3$  соответственно,  $q_1, q_2, q_3$  — вероятности сдвига в отрицательном направлении вдоль осей  $x_1, x_2, x_3$  соответственно,  $p_{12}, p_{13}, p_{23}, p_{123}$  — вероятности сдвига в положительном направлении вдоль осей  $x_1$  и  $x_2$ ,

$x_1$  и  $x_3$ ,  $x_2$  и  $x_3$ ,  $x_1$ ,  $x_2$  и  $x_3$ , соответственно,  $q_{12}$ ,  $q_{13}$ ,  $q_{23}$ ,  $q_{123}$  — вероятности сдвига в отрицательном направлении вдоль осей  $x_1$  и  $x_2$ ,  $x_1$  и  $x_3$ ,  $x_2$  и  $x_3$ ,  $x_1$ ,  $x_2$  и  $x_3$ , соответственно (см. Рис. 3).



**Рис. 3.** Область блуждания для последовательного представления трех стеков в памяти одного уровня в общем случае.

Плоскости  $x_1 + x_2 = m_1 + 1$  и  $x_3 = m_2 + 1$  будут выступать в качестве поглощающих экранов, что соответствует переполнению отрезков памяти для двух и одного стека соответственно.

Плоскости  $x_1 = -1$ ,  $x_2 = -1$  и  $x_3 = -1$  будут отражающими экранами.

Пусть рассматриваемое в качестве математической модели случайное блуждание имеет  $n$  невозвратных состояний. Тогда поглощающие состояния этого случайного блуждания нумеровать не будем, а его невозвратные состояния пронумеруем от 1 до  $n$ , т. е. каждому невозвратному состоянию будет взаимно однозначно соответствовать некоторый номер  $i = 1, \dots, n$ .

Каждое невозвратное состояние этого случайного блуждания с номером  $i$ , где  $i = 1, \dots, n$ , соответствует целочисленной точке в пространстве внутри треугольной призмы с координатами  $(i_1, i_2, i_3)$ , где  $i_1, i_2, i_3 \geq 0$ ,  $i_1 + i_2 \leq m_1$  и  $i_3 \leq m_2$ . В дальнейшем это соответствие будем обозначать  $i = (i_1, i_2, i_3)$ .

Выбор нумерации состояний существенен, так как он влияет на вид матрицы переходных вероятностей. Выберем нумерацию невозвратных состояний рассматриваемого случайного блуждания следующим образом. Зафиксируем  $x_3 = m_2$ , тогда областью блуждания станет внутренность треугольника с вершинами в точках:  $(-1, -1, m_2)$ ,  $(-1, m_1 + 1, m_2)$ ,  $(m_1 + 1, -1, m_2)$ . Пронумеруем невозвратные состояния, начиная с точки  $(m_1, 0, 0)$  снизу вверх справа налево. Пример нумерации состояний (см. Рис. 4).

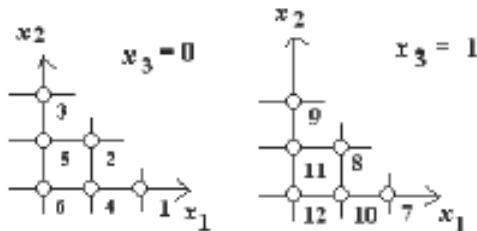


Рис. 4. Пример нумерации состояний.

Перейдем к  $x_3 = m_2 - 1$ , тогда областью блуждания станет внутренняя область треугольника с вершинами в точках:  $(-1, -1, m_2 - 1)$ ,  $(-1, m_1 + 1, m_2 - 1)$ ,  $(m_1 + 1, -1, m_2 - 1)$ . Продолжаем нумерацию аналогичным образом, начиная с точки  $(m_1, 0, 1)$ . И так далее до треугольника с вершинами в точках  $(-1, -1, 0)$ ,  $(-1, m_1 + 1, 0)$ ,  $(m_1 + 1, -1, 0)$ .

**Теорема 1.** При выбранной нумерации состояний подматрица  $Q$  матрицы переходных вероятностей данной цепи, соответствующая переходам из невозвратных состояний в невозвратные, имеет вид:

$$Q_{n \times n} = \begin{pmatrix} W^{m_2+1} & V^{m_2} & 0 & \dots & 0 & 0 \\ U^{m_2} & W^{m_2} & V^{m_2-1} & \dots & 0 & 0 \\ 0 & U^{m_2-1} & W^{m_2-1} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & W^2 & V^1 \\ 0 & 0 & 0 & \dots & U^1 & W^1 \end{pmatrix},$$

$$\text{где } n = n_1(m_2 + 1), \quad n_1 = \frac{(m_1 + 1)(m_1 + 2)}{2};$$

$$U_{n_1 \times n_1}^{l_2} = \begin{pmatrix} K^{m_1+1} & L^{m_1} & M^{m_1-1} & 0 & \dots & 0 & 0 & 0 \\ 0 & K^{m_1} & L^{m_1-1} & M^{m_1-2} & \dots & 0 & 0 & 0 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & K^3 & L^2 & M^1 \\ 0 & 0 & 0 & 0 & \dots & 0 & K^2 & L^1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & K^1 \end{pmatrix},$$

*rge*  $l_2 = 1, \dots, m_2$ ;

$$K_{l_1 \times l_1}^{l_1} = \begin{pmatrix} q_3 & 0 & \dots & 0 & 0 \\ 0 & q_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_3 & 0 \\ 0 & 0 & \dots & 0 & q_3 \end{pmatrix}, \text{ rge } l_1 = 2, \dots, m_1 + 1;$$

$$L_{(l_1+1) \times l_1}^{l_1} = \begin{pmatrix} q_{13} & 0 & \dots & 0 & 0 \\ q_{23} & q_{13} & \dots & 0 & 0 \\ 0 & q_{23} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_{13} & 0 \\ 0 & 0 & \dots & q_{23} & q_{13} \\ 0 & 0 & \dots & 0 & q_{23} \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$M_{(l_1+2) \times l_1}^{l_1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ q_{123} & 0 & \dots & 0 & 0 \\ 0 & q_{123} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_{123} & 0 \\ 0 & 0 & \dots & 0 & q_{123} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$V_{n_1 \times n_1}^{l_2} = \begin{pmatrix} C^{m_1+1} & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ D^{m_1} & C^{m_1} & 0 & 0 & \dots & 0 & 0 & 0 \\ E^{m_1-1} & D^{m_1-1} & C^{m_1-1} & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & C^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & D^2 & C^2 & 0 \\ 0 & 0 & 0 & 0 & \dots & E^1 & D^1 & C^1 \end{pmatrix},$$

*rge*  $l_2 = 1, \dots, m_2$ ;

$$C_{l_1 \times l_1}^{l_1} = \begin{pmatrix} p_3 & 0 & \dots & 0 & 0 \\ 0 & p_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_3 & 0 \\ 0 & 0 & \dots & 0 & p_3 \end{pmatrix}, \text{ rge } l_1 = 2, \dots, m_1 + 1;$$

$$D_{l_1 \times (l_1+1)}^{l_1} = \begin{pmatrix} p_{13} & p_{23} & 0 & \dots & 0 & 0 & 0 \\ 0 & p_{13} & p_{23} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p_{13} & p_{23} & 0 \\ 0 & 0 & 0 & \dots & 0 & p_{13} & p_{23} \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$E_{(l_1+2) \times l_1}^{l_1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ p_{123} & 0 & \dots & 0 & 0 \\ 0 & p_{123} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_{123} & 0 \\ 0 & 0 & \dots & 0 & p_{123} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$W_{n_1 \times n_1}^{l_2} = \begin{pmatrix} H^{m_1+1} & T^{m_1} & A^{m_1-1} & 0 & \dots & 0 & 0 & 0 \\ S^{m_1} & H^{m_1} & T^{m_1-1} & 0 & \dots & 0 & 0 & 0 \\ B^{m_1-1} & S^{m_1-1} & H^{m_1-1} & T^{m_1-2} & \dots & 0 & 0 & 0 \\ 0 & B^{m_1-2} & S^{m_1-2} & H^{m_1-2} & \dots & 0 & 0 & 0 \\ \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & H^3 & T^2 & A^1 \\ 0 & 0 & 0 & 0 & \dots & S^2 & H^2 & T^1 \\ 0 & 0 & 0 & 0 & \dots & B^1 & S^1 & H^1 \end{pmatrix},$$

*rge*  $l_2 = 2, \dots, m_2 + 1$ ;

$$S_{l_1 \times (l_1+1)}^{l_1} = \begin{pmatrix} p_1 & p_2 & 0 & \dots & 0 & 0 & 0 \\ 0 & p_1 & p_2 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p_1 & p_2 & 0 \\ 0 & 0 & 0 & \dots & 0 & p_1 & p_2 \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$B_{l_1 \times (l_1+2)}^{l_1} = \begin{pmatrix} 0 & p_{12} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & p_{12} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p_{12} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & p_{12} & 0 \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$T_{(l_1+1) \times l_1}^{l_1} = \begin{pmatrix} q_1 & 0 & \dots & 0 & 0 \\ q_2 & q_1 & \dots & 0 & 0 \\ 0 & q_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_1 & 0 \\ 0 & 0 & \dots & q_2 & q_1 \\ 0 & 0 & \dots & 0 & q_2 \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$A_{(l_1+2) \times l_1}^{l_1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ q_{12} & 0 & \dots & 0 & 0 \\ 0 & q_{12} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_{12} & 0 \\ 0 & 0 & \dots & 0 & q_{12} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}, \text{ rge } l_1 = 1, \dots, m_1;$$

$$H_{l_1 \times l_1}^{l_1} = \begin{pmatrix} q_2 + q_{12} + q_{23} + q_{123} & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & q_1 + q_{12} + q_{13} + q_{123} \end{pmatrix},$$

$$\text{rge } l_1 = 2, \dots, m_1 + 1;$$

$$H_{1 \times 1}^1 = ( q_1 + q_2 + q_{12} + q_{13} + q_{23} + q_{123} );$$

$$W_{n_1 \times n_1}^1 = \begin{pmatrix} F^{m_1+1} & T^{m_1} & 0 & \dots & 0 & 0 \\ S^{m_1} & F^{m_1} & T^{m_1-1} & \dots & 0 & 0 \\ 0 & S^{m_1-1} & F^{m_1-1} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & F^2 & T^1 \\ 0 & 0 & 0 & \dots & S^1 & F^1 \end{pmatrix};$$

$$F_{l_1 \times l_1}^{l_1} = \begin{pmatrix} q_2 + q_3 + q_{12} + q_{23} + q_{123} & 0 & \dots & 0 & 0 \\ 0 & q_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_3 & 0 \\ 0 & 0 & \dots & 0 & q_1 + q_3 + q_{12} + q_{13} + q_{123} \end{pmatrix},$$

где  $l_1 = 2, \dots, m_1 + 1$ ;

$$F_{1 \times 1}^1 = (q_1 + q_2 + q_3 + q_{12} + q_{13} + q_{23} + q_{123}).$$

Можно найти фундаментальную матрицу  $N = (I - Q)^{-1}$  и вычислить среднее время блуждания [4].

Была написана программа, которая реализует данный алгоритм.

Результаты численных экспериментов приводятся в таблицах 3-4.

Условные обозначения в таблицах 3-4:  $W_1$  – среднее время в случае последовательного представления трех стеков, когда первый и второй стеки объединены в пару, а третий стек расположен отдельно;  $W_2$  – среднее время в случае последовательного представления трех стеков, когда первый и третий стеки объединены в пару, а второй стек расположен отдельно;  $W_3$  – среднее время в случае последовательного представления трех стеков, когда второй и третий стеки объединены в пару, а первый стек расположен отдельно  $W_4$  – среднее время в случае связанного представления трех стеков.

**Таблица 3.** Среднее время в случае последовательного представления трех стеков в памяти одного уровня в случае параллельного выполнения операций для  $p_1 = 0.35, p_2 = p_3 = p_{12} = p_{23} = p_{13} = p_{123} = 0.05, q_1 = q_2 = q_3 = q_{12} = q_{13} = q_{23} = q_{123} = 0.05$ .

$m$	$W_1$		$W_2$		$W_3$		$W_4$
	$m_1$	$T$	$m_1$	$T$	$m_1$	$T$	$T$
10	7	11.15	7	11.15	4	10.78	8.56
20	15	20.68	15	20.68	8	19.88	17.96

Из таблицы 3 видно, что если все вероятности равны, кроме вероятности включения в один из стеков (например, первый), и при этом эта вероятность больше всех остальных вероятностей, то среднее время до переполнения будет больше в том случае, если этот стек поместить в пару.

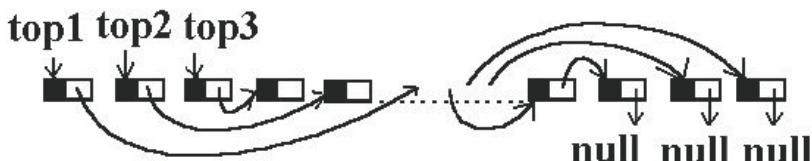
**Таблица 4.** Среднее время в случае последовательного представления трех стеков в памяти одного уровня в случае параллельного выполнения операций для  $p_1 = p_2 = p_3 = 0.01, p_{13} = p_{12} = p_{23} = p_{123} = 0.22, q_1 = q_2 = q_3 = q_{12} = q_{13} = q_{23} = q_{123} = 0.01$ .

$m$	$W_1$		$W_2$		$W_3$		$W_4$
	$m_1$	$T$	$m_1$	$T$	$m_1$	$T$	$T$
10	6	5.29	6	5.29	6	5.29	3.3
20	13	10.10	13	10.10	13	10.10	5.89

Результаты таблицы 4 показывают, что если вероятности  $p_{13}, p_{12}, p_{23}, p_{123}$ , одинаковы, то среднее время до переполнения для всех трех случаев одинаково.

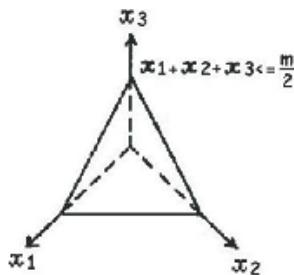
## 2. Связанное представление

Пусть мы работаем в памяти размера  $m$  единиц с тремя стеками, расположенными связно в случае параллельного выполнения операций (см. Рис. 5).



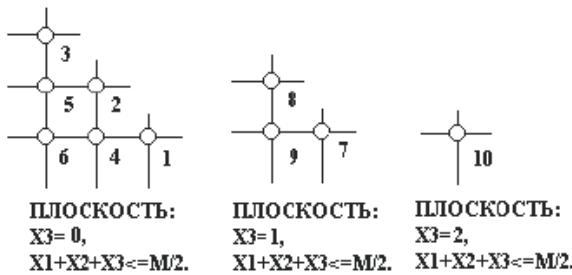
**Рис. 5.** Связанное представление трех стеков в памяти одного уровня в случае параллельного выполнения операций.

В качестве математической модели имеем случайное блуждание в пирамиде:  $x_1 + x_2 + x_3 \leq \frac{m}{2}$  (см. Рис. 6), т.к. половина памяти тратится на связи.



**Рис. 6.** Область блуждания для связанных представления трех стеков в памяти одного уровня в случае параллельного выполнения операций.

Пронумеруем состояния: начиная с плоскости  $x_3 = 0$ ,  $x_1 + x_2 + x_3 \leq \frac{m}{2}$ ,  $x_3 = 1$ ,  $x_1 + x_2 + x_3 \leq \frac{m}{2}$  и т.д. наверх, в каждой из этих плоскостей нумерация состояний начинается вдоль прямой  $x_1 + x_2 = \frac{m}{2} - x_3$  от оси  $x_1$  к оси  $x_2$ ,  $x_1 + x_2 = \frac{m}{2} - x_3 - 1$  и т.д. (см. Рис. 7).



**Рис. 7.** Пример нумерации состояний для  $m = 4$ .

Тогда можно построить матрицу  $Q$  переходных вероятностей однородной поглощающей цепи Маркова.

**Теорема 2.** При выбранной нумерации состояний подматрица  $Q$  матрицы переходных вероятностей данной цепи, соответствующая переходам из невозвратных состояний в невозвратные, имеет вид:

$$Q = \begin{pmatrix} A_{\frac{n*(n+1)}{2}}^{n+1} & B_{\frac{n*(n+1)}{2}} & 0 & \dots & 0 & 0 \\ C_{\frac{(n-1)*n}{2}} & A_{\frac{(n-1)*n}{2}}^n & B_{\frac{(n-1)*n}{2}} & \dots & 0 & 0 \\ 0 & C_{\frac{(n-2)*(n-1)}{2}} & A_{\frac{(n-2)*(n-1)}{2}}^{n-1} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & A_3^2 & B_3 \\ 0 & 0 & 0 & \dots & C_1 & A_1^1 \end{pmatrix},$$

*rge*  $n = \frac{m}{2}$ ,

подматрица  $A_{\frac{n*(n+1)}{2}}^l$  - квадратная матрица размерности  $\frac{n*(n+1)}{2} \times \frac{n*(n+1)}{2}$  имеет вид:

$$A_{\frac{n*(n+1)}{2}}^l =$$

$$\begin{pmatrix} D_{n \times n} & E_{n \times (n-1)} & G_{n \times (n-2)} & \dots & 0 & 0 & 0 \\ F_{n \times (n-1)} & D_{(n-1) \times (n-1)} & E_{(n-1) \times (n-2)} & \dots & 0 & 0 & 0 \\ X_{n \times (n-2)} & F_{(n-1) \times (n-2)} & D_{(n-2) \times (n-2)} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & D_{3 \times 3} & E_{3 \times 2} & G_{3 \times 1} \\ 0 & 0 & 0 & \dots & F_{3 \times 2} & D_{2 \times 2} & E_{2 \times 1} \\ 0 & 0 & 0 & \dots & X_{3 \times 1} & F_{2 \times 1} & D_{1 \times 1} \end{pmatrix},$$

*rge*  $l = 1, \dots, n$ ,

$$D_{n \times n} =$$

$$\begin{pmatrix} q_2 + q_{12} + q_{13} + q_{23} + q_{123} & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & q_1 + q_{12} + q_{13} + q_{23} + q_{123} \end{pmatrix},$$

*rge*  $D_{1 \times 1} = (q_1 + q_2 + q_{12} + q_{13} + q_{23} + q_{123})$ ,

$$A_{\frac{n*(n+1)}{2}}^{n+1} =$$

$$\begin{pmatrix} H_{n \times n} & E_{n \times (n-1)} & G_{n \times (n-2)} & \cdots & 0 & 0 & 0 \\ F_{n \times (n-1)} & H_{(n-1) \times (n-1)} & E_{(n-1) \times (n-2)} & \cdots & 0 & 0 & 0 \\ X_{n \times (n-2)} & F_{(n-1) \times (n-2)} & H_{(n-2) \times (n-2)} & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & H_{3 \times 3} & E_{3 \times 2} & G_{3 \times 1} \\ 0 & 0 & 0 & \cdots & F_{3 \times 2} & H_{2 \times 2} & E_{2 \times 1} \\ 0 & 0 & 0 & \cdots & X_{3 \times 1} & F_{2 \times 1} & H_{1 \times 1} \end{pmatrix},$$

$$H_{n \times n} = \begin{pmatrix} q_2 + q_{12} + t & 0 & 0 & \cdots & 0 & 0 \\ 0 & t & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & t & 0 \\ 0 & 0 & 0 & \cdots & 0 & q_1 + q_{12} + t \end{pmatrix},$$

$$rge\ t = q_3 + q_{13} + q_{23} + q_{123},$$

$$rge\ H_{1 \times 1} = ( q_1 + q_2 + q_3 + q_{12} + q_{13} + q_{23} + q_{123} ),$$

$$E_{n \times (n-1)} = \begin{pmatrix} q_1 & 0 & 0 & \cdots & 0 & 0 \\ q_2 & q_1 & 0 & \cdots & 0 & 0 \\ 0 & q_1 & q_2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & q_2 & q_1 \\ 0 & 0 & 0 & \cdots & 0 & q_2 \end{pmatrix},$$

$$F_{n \times (n-1)} = \begin{pmatrix} p_1 & p_2 & 0 & \cdots & 0 & 0 \\ 0 & p_1 & p_2 & \cdots & 0 & 0 \\ 0 & 0 & p_1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & p_2 & 0 \\ 0 & 0 & 0 & \cdots & p_1 & p_2 \end{pmatrix},$$

подматрица  $B_{\frac{n*(n+1)}{2}}$  - матрица размерности  $\frac{n*(n+1)}{2} \times \frac{(n-1)*n}{2}$  имеет вид:

$$B_{\frac{n*(n+1)}{2}} =$$

$$\left( \begin{array}{ccccccc} 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ T_{n \times n} & 0 & 0 & \dots & 0 & 0 & 0 \\ Z_{(n-1) \times (n)} & T_{(n-1) \times (n-1)} & 0 & \dots & 0 & 0 & 0 \\ V_{(n-2) \times (n)} & Z_{(n-2) \times (n-1)} & T_{(n-2) \times (n-2)} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & T_{3 \times 3} & 0 & 0 \\ 0 & 0 & 0 & \dots & Z_{2 \times 3} & T_{2 \times 2} & 0 \\ 0 & 0 & 0 & \dots & V_{1 \times 3} & Z_{1 \times 2} & T_{1 \times 1} \end{array} \right),$$

$$T_{n \times n} = \begin{pmatrix} p_3 & 0 & \dots & 0 & 0 \\ 0 & p_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_3 & 0 \\ 0 & 0 & \dots & 0 & p_3 \end{pmatrix},$$

$$Z_{(n-1) \times (n)} = \begin{pmatrix} p_{13} & p_{23} & 0 & \dots & 0 & 0 & 0 \\ 0 & p_{13} & p_{23} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p_{13} & p_{23} & 0 \\ 0 & 0 & 0 & \dots & 0 & p_{13} & p_{23} \end{pmatrix},$$

$$V_{(n-2) \times (n)} = \begin{pmatrix} 0 & p_{123} & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & p_{123} & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p_{123} & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & p_{123} & 0 \end{pmatrix},$$

подматрица  $C_{\frac{(n-1)*n}{2}}$  - матрица размерности  $\frac{(n-1)*n}{2} \times \frac{n*(n+1)}{2}$  имеет вид:

$$C_{\frac{(n-1)*n}{2}} =$$

$$\begin{pmatrix} 0 & \dots & 0 & W_{n \times n} & Y_{n \times (n-1)} & J_{n \times (n-2)} & \dots & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & W_{(n-1) \times (n-1)} & Y_{(n-1) \times (n-2)} & \dots & 0 & 0 & 0 \\ \dots & \dots \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & W_{3 \times 3} & Y_{3 \times 2} & J_{3 \times 1} \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & W_{2 \times 2} & Y_{2 \times 1} \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 & W_{1 \times 1} \end{pmatrix},$$

$$W_{n \times n} = \begin{pmatrix} q_3 & 0 & \dots & 0 & 0 \\ 0 & q_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_3 & 0 \\ 0 & 0 & \dots & 0 & q_3 \end{pmatrix},$$

$$Y_{n \times (n-1)} = \begin{pmatrix} q_{13} & 0 & \dots & 0 & 0 \\ q_{23} & q_{13} & \dots & 0 & 0 \\ 0 & q_{23} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q_{13} & 0 \\ 0 & 0 & \dots & q_{23} & q_{13} \\ 0 & 0 & \dots & 0 & q_{23} \end{pmatrix},$$

$$J_{n \times (n-2)} = \begin{pmatrix} 0 & \dots & 0 & 0 & 0 \\ p_{123} & 0 & \dots & 0 & 0 \\ 0 & p_{123} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & p_{123} & 0 \\ 0 & 0 & \dots & 0 & p_{123} \\ 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Построим фундаментальную матрицу  $N = (I - Q)^{-1}$ . Можно найти среднее время блуждания.

Была написана программа, которая реализует данный алгоритм, результаты представлены в таблицах 3-4.

Из таблиц 3-4 видно, что в большинстве случаев среднее время до переполнения больше в случае последовательного представления стеков, чем в случае связанного представления, но не

всегда. Заметим, что в случае, когда все вероятности равны, среднее время до переполнения в случае последовательного представления трех стеков в памяти одного уровня в случае параллельного выполнения операций больше, чем среднее время в случае связанного представления.

### 3. Случай, когда длина информационной части больше единицы

Предположим, что информационная часть имеет произвольный размер  $l$ , тогда сравним два способа связанное и последовательное представления стеков.

Связанное представление: поле связи также имеет единичную длину, размер памяти  $m$  кратен  $l + 1$ . Тогда в качестве математической модели имеем случайное блуждание по целочисленной решетке, с расстоянием между узлами  $l + 1$ , в пирамиде:  $x_1 + x_2 + x_3 \leq \frac{m}{l+1}$ , предполагается, что  $m$  кратно  $l + 1$ .

Пронумеруем состояния также, как и в задаче о связанном представлении трех стеков в случае параллельного выполнения операций (по плоскостям пирамиды). Построим матрицу переходных вероятностей для однородной поглощающей цепи Маркова. Матрица  $Q$  будет иметь такую же структуру, как матрица  $Q$  для случая связанной организации памяти.

Найдем фундаментальную матрицу  $N = (I - Q)^{-1}$ . Вычислим среднее время блуждания.

Далее рассмотрим последовательное представление трех стеков в случае разной длины информационной части, т.е.  $l > 1$ . В этом случае размер памяти  $m$  должен быть кратен  $l$ .

Пронумеруем состояния также, как и в задаче о последовательном представлении трех стеков в случае параллельного выполнения операций (по плоскостям пирамиды). Построим матрицу переходных вероятностей для однородной поглощающей цепи Маркова. Матрица  $Q$  будет иметь такую же структуру, как матрица  $Q$  для случая последовательного представления трех стеков. Вычислим среднее время.

Результаты сравнения этих двух способов представлены в таблицах 5-6.

**Таблица 5.** Сравнение связанного и последовательного представления трех стеков в памяти одного уровня в случае параллельного выполнения операции и размере информационной части большем единицы,  $m = 30$ ,  $p_1 = 0.87$ ,  $p_2 = p_3 = p_{12} = p_{13} = p_{23} = p_{123} = 0.01$ ,  $q_1 = q_2 = q_3 = q_{12} = q_{13} = q_{23} = q_{123} = 0.01$ .

	$l = 2$	$l = 5$
Связанное	11.54	6.27
Последовательное	14.66	6.26

**Таблица 6.** Сравнение связанного и последовательного представления трех стеков в памяти одного уровня в случае параллельного выполнения операции и размере информационной части большем единицы,  $m = 30$ ,  $p_1 = p_2 = p_3 = p_{12} = p_{13} = p_{23} = p_{123} = 0.01$ ,  $q_1 = 0.87$ ,  $q_2 = q_3 = q_{12} = q_{13} = q_{23} = q_{123} = 0.01$ .

	$l = 2$	$l = 5$
Связанное	212.97	93.98
Последовательное	93.37	61.63

Из таблиц 5-6 видно, что с увеличением размера  $l$  для большинства вероятностей среднее время до переполнения в случае связанного представления трех параллельных стеков лучше, чем в случае последовательного представления.

#### 4. СТРАНИЧНОЕ ПРЕДСТАВЛЕНИЕ

Связанное представление структур данных в памяти является частным случаем страничного представления структур данных. В случае страничного представления память разделена на страницы одинакового размера. На связь тратится один элемент каждой страницы, поэтому потери на связи в случае страничного представления стеков будут меньше, чем в случае связанного представления.

Мы предполагаем, что существует механизм работы со списком свободных страниц, который позволяет при заполнении страницы предоставить стеку новую страницу. Если же вершины стека освободили страницу, то эта страница возвращается в список свободных страниц.

Пусть в страничной памяти размера  $m$  единиц (см. Рис. 9) мы работаем с тремя стеками в случае параллельного выполнения операций.



**Рис. 9.** Страницное представление трех стеков в памяти одного уровня в случае параллельного выполнения операций.

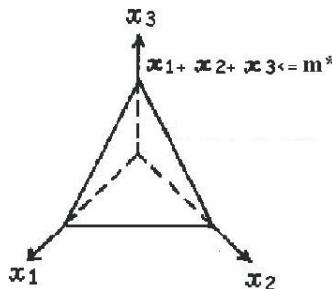
Сначала определим оптимальный размер страницы. Пусть размер страницы  $x$  единиц памяти. Тогда количество страниц  $\frac{m}{x}$ . Обозначим через  $M(x)$  количество единиц памяти, которое тратится непосредственно под данные. Будем считать, что  $M(x) = m - \frac{m}{x} - \frac{2*(x-1)}{2}$ , т.к.  $\frac{m}{x}$  единиц памяти тратятся на связи; также после переполнения одного из стеков, в том случае, когда список свободной памяти уже пуст, работа заканчивается, мы предполагаем, что две другие страницы, с которыми работали два других стека, заполнены не полностью данными, предположим, что на половину, т.е.  $\frac{2*(x-1)}{2}$  единиц.

Тогда  $M'(x) = \frac{m}{x^2} - 1$ , т.е. оптимальный размер страницы  $x = \sqrt{m}$ .

Если проверить знак второй производной, то можно убедиться, что эта точка максимума. Мы работаем с дискретной моделью, поэтому необходимо проверить два значения - округлить  $x$  до целого снизу и сверху.

В качестве математической модели имеем случайное блуждание в пирамиде:  $x_1 + x_2 + x_3 \leq m^*$ , где  $m^* = m - \frac{m}{x}$  для случая, когда память используется полностью (максимально), т.е. после переполнения из одного из стеков, когда в списке свободной памяти нет свободных страниц, две другие страницы, с которыми работали два других стека, тоже заполнены полностью, и  $m^* = m - \frac{m}{x} - 2 * (x - 2)$  для случая, когда память используется не полностью (минимально), т.е. после переполнения из одного из стеков, когда в списке свободной памяти нет свободных страниц,

две другие страницы, с которыми работали два других стека, заполнены минимально, в них содержится только по одной единице данных (см. Рис.10).



**Рис. 10.** Область блуждания для страничного представления трех стеков в памяти одного уровня в случае параллельного выполнения операций.

Пронумеруем состояния также, как и в задаче о связанным представлении трех стеков в случае параллельного выполнения операций (по плоскостям пирамиды). Построим матрицу переходных вероятностей для однородной поглощающей цепи Маркова. Матрица  $Q$  для случая страничной организации памяти будет иметь такую же структуру, как матрица  $Q$  для случая связной организации памяти.

Построим фундаментальную матрицу  $N = (I - Q)^{-1}$ . Вычислим две оценки для среднего времени (максимальную и минимальную), обозначим их  $T_{max}$  и  $T_{min}$ .

Была написана программа, которая реализует данный алгоритм. Результаты представлены в таблицах 7-8, там же, для сравнения приводится среднее время до переполнения в случае связного представления трех стеков.

**Таблица 7.** Страницное и связное представление трех стеков в памяти одного уровня в случае параллельного выполнения операции  $p_1 = 0.87, p_2 = p_3 = p_{12} = p_{13} = p_{23} = p_{123} = 0.01, q_1 = q_2 = q_3 = q_{12} = q_{13} = q_{23} = q_{123} = 0.01$ .

$m$	$x$	$T_{min}$	$T_{max}$
16	4	9.42	13.66
25	5	15.79	22.23

$m$	$T$
16	9.42
26	14.72

**Таблица 8.** Страницное и связанное представление трех стеков в памяти одного уровня в случае параллельного выполнения операции  $p_1 = p_2 = p_3 = p_{12} = p_{13} = p_{23} = p_{123} = 0.01$ ,  $q_1 = 0.87$ ,  $q_2 = q_3 = q_{12} = q_{13} = q_{23} = q_{123} = 0.01$ .

$m$	$x$	$T_{min}$	$T_{max}$
16	4	163.64	263.51
25	5	314.77	470.45

$m$	$T$
16	163.64
26	289.08

Из таблиц 7-8 видно, что среднее время до переполнения в случае страницного представления трех стеков в памяти одного уровня в случае параллельного выполнения операций больше, чем среднее время в случае связанного представления таких же стеков.

С увеличением размера памяти  $m$  ( $m > 16$ ) среднее время для связанного представления трех стеков будет хуже, чем минимальная оценка для среднего времени в случае страницного представления трех стеков. Проверим, с какого  $m$  среднее время в случае связанного представления будет всегда меньше, чем минимальная оценка для среднего времени в случае страницного представления трех стеков. Для этого необходимо сравнить области блуждания. Как только область блуждания будет больше, так среднее время будет больше:

$$m - \frac{m}{x} - 2 * (x - 2) > m - \frac{m}{2}, x = \sqrt{m} =>$$

$$m \subset (-\infty, 4) \bigcup (16, +\infty),$$

т.е. начиная с  $x = \sqrt{17}$  ( $x = 4$  или  $x = 5$ ) и больших, минимальная оценка для среднего времени в случае страницного представления трех стеков всегда лучше, чем среднее время в случае связанного представления.

Если в качестве размера страницы рассматривается не  $x = \sqrt{m}$ , а любой другой произвольный размер, то это утверждение также верно:

$$m - \frac{m}{x} - 2 * x + 4 > \frac{m}{2}, m * (x - 2) > 4 * x * (x - 2),$$

для  $x > 2$  минимальная оценка для среднего времени всегда лучше, чем среднее время в случае связного представления, если  $x < \frac{m}{4}$ , ( $x \leq 2$  - для нашей задачи не имеет смысла).

## 5. ЗАКЛЮЧЕНИЕ

В данной работе рассмотрено несколько способов представления трех стеков в памяти одного уровня в случае параллельного выполнения операций. Были предложены математические модели и получены численные результаты для оценки этих методов с точки зрения максимизации среднего времени работы до переполнения памяти.

Среднее время в случае последовательного представления, для большинства наборов вероятностей, лучше, чем среднее время во всех остальных способах представления.

Среднее время в случае страничного представления трех стеков всегда больше, чем среднее время в случае связанного представления.

## Литература

1. Кнут Д. Искусство программирования для ЭВМ. М.: Мир, 1976. Т.1.
2. Koopman P. Stack Computers. Ellis Horwood, 1989. URL: [http://www.cs.cmu.edu/~koopman/stack\\_computers/](http://www.cs.cmu.edu/~koopman/stack_computers/)
3. В. Феллер. Введение в теорию вероятностей и ее приложения. Мир, Москва, 1964.
4. Кемени Дж., Снелл Дж. Конечные цепи Маркова. М:Наука, 1970.

УДК 519.711

**ЗАДАЧА РАЦИОНАЛЬНОГО РАЗМЕЩЕНИЯ  
ССЫЛОК В РЕГЛАМЕНТИРУЕМОЙ ЛОКАЛИЗОВАННОЙ  
СИСТЕМЕ ИНТЕРНЕТ-РЕСУРСОВ**

А. А. Печников

В статье рассматривается задача рационального размещения взаимных ссылок на Интернет-ресурсы, входящие в некоторую локализованную систему. Рассмотрен случай системы, в которой каждый участник обязан поставить определенное количество ссылок на других участников с целью увеличения их значимости таким образом, чтобы итоговые значения значимости как можно меньше отличались от среднего значения по всем участникам системы (случай "социального" решения).

**1. ВВЕДЕНИЕ**

Под интернет-ресурсом будем понимать информацию, размещенную на некотором www-сервере и имеющую собственный URL в протоколе http.

Важной характеристикой интернет-ресурса является его ссылочная популярность, благодаря которой ресурс успешно (или, наоборот, неуспешно) индексируется поисковыми системами. Реализация алгоритмов ссылочной популярности существенно зависит от поисковой системы - PageRank (в Google), ВИЦ (в Яндексе), ИЦ (в Апорте). Можно сказать, что значимость интернет-ресурса зависит от количества ссылок на данный ресурс с учетом значимости самих ссылающихся интернет-ресурсов, однако конкретные детали учета влияния того или иного показателя являются коммерческой тайной поисковых систем. Наиболее подробные публикации посвящены PageRank, например, [1].

Мы будем использовать термин "значимость интернет-ресурса" в смысле некоторой более общей характеристики, определяемой как интегральная функция различных типов ссылочной популярности (в частном случае это может быть, к примеру, конкретный тип индекса цитирования).

Будем называть интернет-системой некоторое множество интернет-ресурсов с определенными на этом множестве связями в виде прямых ссылок с ресурса на ресурс. (Как правило, ссылки должны находиться на заглавной странице соответствующего интернет-ресурса).

Локализованной системой интернет-ресурсов будем называть некоторое подмножество всех интернет-ресурсов, выделенных по заданному смысловому критерию. Примером локализованной интернет-системы может служить множество всех сайтов органов государственной власти Республики Карелия. Замеры ссылочной популярности многих из этих сайтов дают низкие значения, при этом отсутствуют и прямые ссылки между сайтами (несмотря на то, что они должны формировать единое информационное поле органов власти республики).

Из определения значимости следует, что увеличение количества прямых ссылок на интернет-ресурс ведет к росту его значимости. При этом, очевидно, что количество прямых ссылок, выставляемых на главной странице, ограничено (в более общем случае за ссылки на авторитетных ресурсах еще и необходимо платить). Возникает задача расстановки прямых ссылок на сайтах локализованной интернет-системы при некоторых заданных ограничениях с целью увеличения значимости ресурсов, входящих в состав системы.

Например, в случае сайтов органов власти Республики Карелия целью может быть увеличение значимости каждого сайта таким образом, чтобы итоговые значения значимости как можно меньше отличались от среднего значения по всем участникам системы (случай "социального" решения). В отличие от локализованной системы Интернет-ресурсов некоторого бизнес-сообщества (например, сайтов строительных фирм города Петрозаводска), в системе сайтов органов власти расстановка ссылок может быть определена некоторым правовым актом, поэтому такие системы будем называть регламентированными.

## 2. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДЛЯ СЛУЧАЯ РЕГЛАМЕНТИРУЕМОЙ ЛОКАЛИЗОВАННОЙ СИСТЕМЫ

Рассмотрим регламентируемую локализованную Интернет-систему со следующими характеристиками:

п — количество участников системы,

для каждого участника системы известны:

- $c_i$  - значимость i-го участника системы,
- $m_i$  - количество прямых ссылок i-го участника на других участников.

Определим матрицу ссылок как

$$X = (x_{ij}), \quad i, j = \overrightarrow{1, n},$$

где

$$x_{ij} = \begin{cases} 1, & \text{если существует ссылка от i-го участника к j-му,} \\ 0, & \text{если ссылки не существует.} \end{cases}$$

Сформулируем следующие ограничения:

$$x_{ii} = 0, \quad i = \overrightarrow{1, n}, \quad (1)$$

$$\forall i : \sum_{i=1}^n x_{ij} = m_i \quad (2)$$

Ограничение (1) означает, что участнику нельзя давать ссылку на самого себя, а ограничение (2) - то, что i-м участником должно быть выставлено ровно  $m_i$  исходящих ссылок.

Считаем, что изменение значимости каждого участника при заданной матрице X вычисляется по формуле

$$\tilde{c}_j = c_j + \alpha \cdot \sum_{i=1}^n x_{ij} \cdot c_i \quad (3)$$

Здесь  $\alpha$  - некоторый коэффициент,  $0 < \alpha \leq 1$ .

Содержательно это означает, что при установлении ссылки от i-го участника на j-го, т.е. при  $x_{ij} = 1$ , значимость j-го участника возрастает на некоторую часть значимости i-го участника.

**Замечание :** в общем случае полагаем, ссылки между участниками локальной интернет-системы отсутствуют вообще. Понятно, что если некоторые ссылки есть, то задача будет сводиться к рассматриваемой путем введения дополнительных ограничений типа  $x_{ij} = 1$  для некоторых i,j.

В качестве целевой функции для "социальной" математической модели, возьмем функцию среднеквадратичного отклонения по всем участникам. В этом случае целевая функция оптимизационной задачи имеет следующий вид:

$$F(X) = \sum_{j=1}^n \left( \frac{\sum_{k=1}^n \tilde{c}_k}{n} - \tilde{c}_j \right)^2 \rightarrow \min_{x_{ij}}. \quad (4)$$

### 3. ПРИМЕР

Взяты четыре реальных карельских сайта с их ИЦ (по Яндексу). В качестве коэффициентов значимости сайтов,  $c_i$  и  $m_i$ ,  $i = \overrightarrow{1, 4}$  взяты произвольные, но близкие к возможным значениям.

<http://www.gov.karelia.ru>,  $c_1 = 1800$ ,  $m_1 = 2$ ,

<http://minfin.karelia.ru>,  $c_2 = 40$ ,  $m_2 = 3$ ,

<http://www.mzsocial-rk.ru>,  $c_3 = 40$ ,  $m_3 = 1$ ,

<http://www.petrozavodsk-mo.ru>,  $c_4 = 110$ ,  $m_4 = 2$ .

Оптимальное решение задачи (1)-(4) приведено ниже для  $\alpha = 0, 1$ :

$$\mathbf{X}^* = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Вектор новых значений значимости сайтов равен:

$$\tilde{\mathbf{C}} = (1840 \ 1950 \ 1990 \ 190)$$

Оптимальное значение  $F(X^*) = 1.951E+06$ , тогда как в случае нулевой матрицы X имеем  $F(X) = 2.265E+06$ . При этом значимость каждого участника системы возросла, что положительно скажется в дальнейшем на посещаемости этих сайтов.

### 4. ИССЛЕДОВАНИЕ ЗАДАЧИ

Исследуя более подробно сумму  $\tilde{c}_j$  с учетом ее вида (3) и ограничений (2), имеем:

$$\sum_{j=1}^n \tilde{c}_j = \sum_{j=1}^n c_j + \alpha \cdot \sum_{i=1}^n \sum_{j=1}^n x_{ij} \cdot c_i = \sum_{j=1}^n c_j + \alpha \cdot \sum_{i=1}^n c_i \cdot m_i.$$

Введем обозначения  $S = \sum_{j=1}^n c_j$ ,  $S_m = \sum_{i=1}^n c_i \cdot m_i$  и  $R = (S + \alpha \cdot S_m)/n$ .

Тогда целевую функцию (4) можно записать в виде:

$$F(X) = \sum_{l=1}^n (R - c_l - \alpha \cdot \sum_{r=1}^n x_{rl} \cdot c_r)^2.$$

Дифференцируя  $F(X)$  по  $x_{ij}$  после ряда преобразований получим:

$$\frac{\partial F(X)}{\partial x_{ij}} = 2 \cdot (R - c_j - \alpha \cdot \sum_{r=1}^n x_{rj} \cdot c_r) \cdot (-\alpha \cdot c_i).$$

Приравнивая частные производные по  $x_{ij}$  нулю, получим  $n$  уравнений вида:

$$j = \overrightarrow{1, n} : \sum_{r=1}^n x_{rj} \cdot c_r = R/\alpha - c_j.$$

В систему уравнений входят еще  $n$  уравнений из ограничения (2)

$$i = \overrightarrow{1, n} : \sum_{j=1}^n x_{ij} = m_i.$$

Кроме того, по ограничению (1),  $i = \overrightarrow{1, n} : x_{ii} = 0$ . Таким образом, для нахождения оптимального решения требуется решить систему  $2 \cdot n$  уравнений с  $n^2 - n$  неизвестными.

Составим функцию Лагранжа:

$$L(X, \lambda) = \sum_{j=1}^n \left( \frac{\sum_{k=1}^n c_k}{n} - \tilde{c}_j \right)^2 + \sum_{i=1}^n \lambda_i \cdot (m_i - \sum_{j=1}^n x_{ij}).$$

По теореме Куна-Таккера [2], оптимальное значение  $X^*$  удовлетворяет условию:  $(X^*, \lambda^*)$  - седловая точка функции  $L(X, \lambda)$ , то есть

$$L(X^*, \lambda) \leq L(X^*, \lambda^*) \leq L(X, \lambda^*).$$

При этом в седловой точке должны выполняться следующие условия:

$$\frac{\partial L}{\partial X} |_{X=X^*} \geq 0,$$

$$(\frac{\partial L}{\partial X} |_{X=X^*}, X^*) = 0,$$

$$(\frac{\partial L}{\partial \lambda} |_{\lambda=\lambda^*}, \lambda^*) \leq 0,$$

$$X^* \geq 0, \lambda^* \geq 0.$$

Дифференцируя  $L(X, \lambda)$  по  $x_{ij}$  получим:

$$\frac{\partial L(X, \lambda)}{\partial x_{ij}} = 2 \cdot (R - c_j - \alpha \cdot \sum_{r=1}^n x_{rj} \cdot c_r) \cdot (-\alpha \cdot c_i) - \lambda_i.$$

В матрице

$$\mathbf{X}^* = \begin{pmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1k} & \dots & x_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_{i1} & \dots & x_{ij} & \dots & x_{ik} & \dots & x_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & \dots & x_{nj} & \dots & x_{nk} & \dots & x_{nn} \end{pmatrix}$$

в  $i$ -ой строке для любых  $j, k$ :

$$\text{из } x_{ij} > 0 \Rightarrow \lambda_i = -2 \cdot \alpha \cdot c_i \cdot (R - c_j - \alpha \cdot \sum_{r=1}^n x_{rj} \cdot c_r)$$

$$\text{и из } x_{ik} > 0 \Rightarrow \lambda_i = -2 \cdot \alpha \cdot c_i \cdot (R - c_k - \alpha \cdot \sum_{r=1}^n x_{rk} \cdot c_r).$$

Отсюда для любой пары  $(j, k)$  при  $x_{ij} > 0, x_{ik} > 0$  должно выполняться условие:

$$c_j - c_k = \alpha \cdot \sum_{r=1}^n (x_{rk} - x_{rj}) \cdot c_r. \quad (A1)$$

С другой стороны, для любой пары  $(i, j)$  из

$$x_{ij} = 0 \Rightarrow \lambda_i < -2 \cdot \alpha \cdot c_i \cdot (R - c_j - \alpha \cdot \sum_{r=1}^n x_{rj} \cdot c_r).$$

Отсюда в качестве  $\lambda_i$  выбираем наибольшее значение по  $i$ :

$$\tilde{\lambda}_i = \max_j \{-2 \cdot \alpha \cdot c_i \cdot (R - c_j - \alpha \cdot \sum_{r=1}^n x_{rj} \cdot c_r)\}.$$

Обозначим  $\mu_i = \tilde{\lambda}_i / (2 \cdot \alpha \cdot c_i) + R$  и получим условие

$$\mu_i = \max_j \{c_j + \alpha \cdot \sum_{r=1}^n x_{rj} \cdot c_r\}. \quad (A2)$$

## 5. Алгоритмы

Условия A1, A2 содержательно можно трактовать следующим образом: при принятии решения о том, устанавливать или нет ссылку от i-го участника к j-му выбирается участник с наименьшим значением  $\mu_i$  на данном шаге, который при этом получает новое значение  $c_j$ , наиболее близкое к среднему по столбцу.

На этой идеи основан реализованный авторами вариант "жадного" алгоритма. Сравнение данного алгоритма с алгоритмом полного перебора на задачах небольших размерностей дало незначительные отклонения от оптимального решения.

Таблица 1

№	n	$\alpha$	$m_1$	$m_2$	$m_3$	$m_4$	$m_5$	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	опт	прибл
1	4	1	2	3	1	2		1000	40	40	110		66,1E4	77,1E4
2	4	0,2	2	3	1	2		1000	40	40	110		47,9E4	48,4E4
3	4	0,1	2	3	1	2		1000	40	40	110		55,8E4	56,0E4
4	5	0,2	2	3	1	2	2	1000	40	40	110	500	39,5E4	43,0E4
5	5	0,1	2	3	1	2	2	1000	40	40	110	500	52,4E4	52,6E4
6	5	0,1	2	2	1	2	2	150	40	60	100	120	2471	2471

Из таблицы можно заметить, что приближенный алгоритм решения задачи дает вполне приемлемую точность найденных решений (отклонение от оптимального от 0,875 до 1,0). При этом с уменьшением значения  $\alpha$  и при невысоком разбросе значимостей точность решения возрастает.

## Литература

1. Craven Ph. "Google's PageRank Explained and how to make the most of it", <http://www.webworkshop.net/pagerank.html>.
2. Васильев Ф.П. Численные методы решения экстремальных задач. -М.: Наука, 1988.

УДК 519.6, 539.2

**РАЗНОСТНАЯ СХЕМА ДЛЯ МОДЕЛИРОВАНИЯ  
ТДС-СПЕКТРА ДЕГИДРИРОВАНИЯ МЕТАЛЛОВ  
С УЧЕТОМ ТЕПЛОПОГЛОЩЕНИЯ И СЖАТИЯ**

Н. И. Родченкова

Моделируется кинетика выделения водорода из порошка гидрида металла с учетом динамики межфазового перехода, сжатия частиц порошка гидрида, десорбционных процессов на поверхности и теплопоглощения на границе раздела фаз. Представлена разностная схема для краевой задачи с подвижными границами применительно к экспериментальному методу термодесорбционной спектрометрии (ТДС). Приведены результаты численного моделирования.

**1. ВВЕДЕНИЕ**

В [1] представлена диффузационная модель с учетом теплопоглощения на границе раздела фаз и эффекта сжатия. Данная публикация является продолжением этой работы. В статье рассматривается численный метод (построение разностной схемы) решения соответствующей краевой задачи с нелинейными динамическими граничными условиями и подвижными границами. Соответствующая библиография имеется в [2–7]. Модель представлена применительно к экспериментальному методу термодесорбционной спектрометрии (ТДС). Образец порошкообразного гидрида металла помещают в камеру с последующим постоянным вакуумированием и относительно медленным нагревом. По регистрируемому с помощью масс-спектрометра десорбционному потоку водорода из образца судят о тех или иных характеристиках

его взаимодействия с исследуемым материалом. Рассмотрим высокотемпературный пик потока, когда диффузия является одним из лимитирующих факторов.

## 2. ОСНОВНЫЕ ПОНЯТИЯ И ОБОЗНАЧЕНИЯ

При распаде рассматриваемого модельного гидрида происходит существенное увеличение плотности (уменьшение объема). Это связано с перестройкой решетки. Для некоторых гидридов эффект достигает нескольких десятков процентов. Пусть  $\gamma$  — коэффициент сжатия: в процессе разложения гидрида объем с  $V$  уменьшается до значения  $\gamma V$  ( $0 < \gamma < 1$ ,  $\gamma = \text{const}$ ,  $T \in [T^-, T^+]$ ). Размеры частиц считаем уже достаточно малыми, чтобы возникающие напряжения не приводили к дальнейшему измельчению. Будем стремиться учесть эффект минимальными математическими средствами. Обозначим  $L(t)$  — радиус сжимающейся сферической частицы порошка гидрида,  $\rho(t)$  — радиус сжимающегося гидридного ядра. Уменьшение  $L(t)$  инициируется «пропавшим» объемом в результате перестройки решетки на границе раздела фаз. Сокращая множитель  $4\pi/3$ , имеем связь  $\rho$  и  $L$  в форме  $(1 - \gamma)[\rho_0^3 - \rho^3] = L_0^3 - L^3$ . Здесь  $L_0 = L(0)$ ,  $\rho_0 = \rho(0)$ , начальный момент времени  $t = 0$  соответствует началу всплеска десорбции. Чтобы существенно не усложнять модель, полагаем, что внешняя корка как-бы «оседает в целом». Пусть  $c(t, r)$  — концентрация атомарного водорода, растворенного в корке металла (в  $\alpha$  фазе) толщины  $(L - \rho)$  в момент времени  $t$ ;  $Q$  — концентрация водорода в гидриде ( $\beta$  фазе);  $b(T)$  — коэффициент десорбции;  $D(T)$  — коэффициент диффузии;  $\mu$  — кинетическая константа ( $\mu \approx 1.46 \cdot 10^{21} \text{ 1/cm}^2 \cdot \text{s} \cdot \text{Torr}$ );  $p(t)$  — давление молекулярного водорода в камере;  $s$  — коэффициент прилипания водорода к поверхности (диссоциации части молекул  $H_2$  на атомы). Нагрев на поверхности обычно линейный:  $T_L(t) \equiv T(t, L) = vt + T_0$ . Считаем, что к начальному моменту времени ( $t = 0$  соответствует началу пика ТДС-спектра) растворенный в гидриде водород покинул образец и  $Q = c_{\beta}^{crit} = \text{const}$  является критической концентрацией. В процессе дальнейшего нагрева происходит распад  $\beta$  фазы и соответствующая активизация десорбции. Плотность десорбционного потока моделируем квадратичной зависимостью от концентрации атомов в приповерхностном слое (вариант объемной

десорбции):  $J(t) = b(T)c^2(t, L)$ . Зависимость параметров от температуры считаем аррениусовской. Например, для коэффициента десорбции  $b(T) = b_0 \exp\{-E_b/[RT]\}$ ,  $T \in [T^-, T^+]$ , где  $E_b$  – энергия активации,  $b_0 = \text{const}$ ,  $R = \text{const}$ . Давление молекулярного водорода в камере инициирует плотность потока  $\mu s(T)p(t)$  атомов  $H$  на поверхность (в приповерхностный объем). Дегидрирование предполагает  $\mu s p < bc_L^2$ ,  $c_L \equiv c(t, L)$ . Лишь часть десорбированного водорода возвращается обратно в результате откачки  $H_2$ .

Модель будем излагать в одночастичном приближении. Для моделирования порошка использовалось дискретное нормальное распределение частиц по радиусам. Это, в некотором смысле, «компенсирует» погрешность сферической аппроксимации. При численном моделировании использовалось 10-15 значений  $L_0$ .

### 3. ЭФФЕКТ ОБЪЕМНОГО СЖАТИЯ

Обе границы подвижны:  $\rho = \rho(t)$ ,  $L = L(t)$ ,  $(1 - \gamma)[\rho_0^3 - \rho^3] = L_0^3 - L^3$ . Плотность потока атомов водорода (по направлению  $\mathbf{r}_0$  вдоль радиуса,  $|\mathbf{r}_0| = 1$ ) моделируем выражением

$$j(t, r) = -D(T) \frac{\partial c}{\partial r} + c(t, r)v(t, r),$$

где  $v(t, r)$  – «встречная скорость среды» ( $v < 0$ ). Уравнение диффузии с учетом сферической симметрии примет вид:

$$\frac{\partial c}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \left[ D(T) \frac{\partial c}{\partial r} - c(t, r)v(t, r) \right] \right).$$

Уточним выражение для  $v(t, r)$ . В момент времени  $t = t_0$  мысленно выделим шар  $V(r)$ ,  $r > \rho$ , и определим скорость изменения его объема в процессе сжатия. Обозначив радиус через  $R(t)$  ( $R(t_0) = r$ ), получим

$$\dot{V}(r) = \frac{d}{dt} \Big|_{t_0} V(R(t)) = 4\pi r^2 \dot{R}(t_0).$$

В указанных предположениях (напряжения некритичны, нет распространения) должно быть  $\dot{V}(r) = \dot{V}(L)$ , т. е.  $r^2 \dot{R} = L^2 \dot{L}$ . Пропадающий объем компенсируется внешними слоями. Поэтому скорость движения среды вдоль радиуса представим в форме

$$v(t, r) = L^2 \dot{L} / r^2 = (1 - \gamma) \rho^2 \dot{\rho} / r^2, \quad r \in [\rho, L].$$

Отдельно обоснуем  $v(t, \rho) = (1 - \gamma)\dot{\rho}$ : это именно та часть скорости  $\dot{\rho}$  движения границы раздела фаз, которая инициируется сжатием.

Обозначим через  $S(r)$  площадь сферы радиуса  $r$ ,  $T_L(t) \equiv T(t, L(t))$ . Дифференцируем по времени соотношение материального баланса:

$$\begin{aligned} QV(\rho) + 4\pi \int_{\rho(t)}^{L(t)} r^2 c(t, r) dr + \int_0^t S(L) b(T_L) c^2(\tau, L) d\tau &= \text{const}, \\ Q\rho^2 \dot{\rho} + \int_\rho^L r^2 c_t(t, r) dr + c(t, L)L^2 \dot{L} - c(t, \rho)\rho^2 \dot{\rho} + b(T_L)c^2(t, L)L^2 &= 0, \\ [Q - \gamma c(t, \rho)]\rho^2 \dot{\rho} - \rho^2 D(T[t, \rho]) \frac{\partial c}{\partial r} \Big|_\rho + b(T_L)c_L^2 L^2 + L^2 D(T_L) \frac{\partial c}{\partial r} \Big|_L &= 0. \end{aligned}$$

Первые два слагаемых относятся к процессам на границе раздела фаз  $r = \rho$ , остальные — на поверхности  $r = L$ . Вследствие баланса потоков на внешней границе  $b(T_L)c^2(t, L) = -D(T_L)c_r(t, L)$  (нет накопления  $H$  на поверхности) приходим к условию Стефана

$$[Q - \gamma c(t, \rho)]\dot{\rho}(t) = D(T[t, \rho]) \frac{\partial c}{\partial r} \Big|_\rho.$$

Уравнение движения границы раздела фаз можно вывести и независимо из баланса

$$Q \delta V - c(t, \rho)\gamma \delta V = -S(\rho)D(T[t, \rho])c_r(t, \rho)\Delta t + o(\Delta t), \quad \delta V = -\Delta V.$$

Тогда уже граничное условие нелинейное третьего рода окажется следствием.

#### 4. МОДЕЛЬ ТЕПЛОПОГЛОЩЕНИЯ

В контексте метода ТДС прогрев частиц гидрида значительно быстрее диффузии. Поэтому разумно искать распределение температуры в растущей корке металла с растворенным дифундирующими водородом в форме квазистационара  $T(t, r) = A^\circ(t) + B^\circ(t)/r$ ,  $r \in [\rho, L]$ . Скорость нагрева мала: распределение температуры не сильно отличается от равномерного, имеется относительно небольшое понижение  $T$  при  $r = \rho$  по сравнению с  $r = L$ . Поэтому  $B^\circ(t) < 0$  (функция  $T(t, r)$  выпукла вверх).

Рассмотрим начальный вариант учета эндотермичности распада гидрида. Полагаем, что в частице порошка за время распада

гидрид не успевает существенно изменить свою температуру в отличие от окружающей корки металла. На малом промежутке времени  $[t, t + \Delta t]$  образуется шаровой слой раствора объема  $\delta V = [\rho^3 - (\rho + \Delta\rho)^3]4\pi/3$ ,  $\Delta\rho < 0$ ,  $\Delta\rho = \rho(t + \Delta t) - \rho(t) \approx \dot{\rho}(t)\Delta t$ . Количество поглощенного тепла  $\theta\delta V$ , пропорциональное этому объему, равно величине  $P4\pi\rho^2\Delta t$  с точностью до  $o(\Delta t)$ . Здесь обозначили  $P = \lambda\partial_r T = -\lambda B^\circ/\rho^2$  — плотность теплового потока. Приравнивая два выражения и деля на  $\Delta t \rightarrow 0$ , определяем коэффициент  $B^\circ$ :  $B^\circ(t) = \kappa\dot{\rho}(t)\rho^2(t)$ ,  $\kappa = \theta/\lambda$ . В пределах пика ТДС-спектра величину  $\kappa$ , характеризующую теплопроводящие свойства металла и энергетику распада, считаем положительной константой. Коэффициент  $A^\circ(t)$  определяется контролируемой температурой  $T_L(t)$  на поверхности. Итак, гиперболическое по  $r$  распределение  $T(t, r)$  несколько уменьшается при  $r \rightarrow \rho + 0$  по сравнению с  $T(t, L)$  из-за эндотермичности распада, а ядро прогрето практически равномерно. Из-за большого различия в теплопроводящей способности (например, гидрид — солеподобный) имеется разрыв температуры на границе раздела фаз. При  $\rho(t) \rightarrow 0$  имеем  $B^\circ(t) \rightarrow 0$  и процесс остаточной дегазации начинается и протекает при равномерном прогреве частицы.

## 5. ЛОКАЛЬНОЕ РАВНОВЕСИЕ И НАЧАЛЬНЫЕ ДАННЫЕ

В равновесной ситуации концентрация атомов  $H$  в металле устанавливается на уровне  $\bar{c} = \eta Q < Q$ . При высокой температуре гидрид распадается достаточно интенсивно и в условиях диффузационного барьера можно считать, что на границе раздела фаз имеет место локальное равновесие  $c(t, \rho) = \bar{c}$ .

Уравнению диффузии  $c_t = D(t)(c_{rr} + 2c_r/r)$  удовлетворяет независящая от времени функция  $\varphi(r) = A + B/r$ . Примем ее за начальное распределение  $c(0, r)$  в слое  $r \in [\rho_0, L_0]$ : До начала интенсивного распада гидрида ( $t = 0$ ) в результате объемной десорбции происходит медленный рост начальной корки металла с равномерным прогревом. Таким образом, когда граница на начальном этапе движется очень медленно, вид распределения растворенного водорода (гиперболический квазистационар) почти не будет меняться со временем. Обозначим  $c(0, L_0) = c_0$ , тогда:  $\varphi(\rho_0) = \bar{c}$ ,

$$\varphi(L_0) = c_0 < \bar{c} \Rightarrow$$

$$\varphi(r) = A + \frac{B}{r} = \frac{L_0 c_0 - \bar{c} \rho_0}{L_0 - \rho_0} + L_0 \rho_0 \cdot \frac{\bar{c} - c_0}{L_0 - \rho_0} \cdot \frac{1}{r}.$$

Параметр  $c_0$  находим из граничного условия на поверхности при  $t = 0$ :

$$b(T_0)c_0^2 = -D(T_0) \left. \frac{\partial \varphi}{\partial r} \right|_{L_0} = D(T_0) \cdot \frac{\bar{c} - c_0}{L_0 - \rho_0} \cdot \frac{\rho_0}{L_0}.$$

Это квадратное уравнение относительно  $c_0$ , корни вещественны и разных знаков, так что приповерхностная концентрация определяется однозначно. Отметим, что начальное распределение  $\varphi(r)$  в тонкой корке металла практически не влияет на модельный ТДС-пик, поскольку весь водород сосредоточен в гидридном ядре.

## 6. КРАЕВАЯ ЗАДАЧА ТДС-ДЕГИДРИРОВАНИЯ

Компактно представим модель:  $c_0 \equiv c(0, L_0)$ ,  $\rho(0) = \rho_0$ ,

$$\frac{\partial c}{\partial t} = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \left[ D(T[t, r]) \frac{\partial c}{\partial r} - c v \right] \right), \quad t \in (0, t_*), \quad r \in (\rho(t), L(t)),$$

$$c(0, r) = A + \frac{B}{r} = \frac{L_0 c_0 - \bar{c} \rho_0}{L_0 - \rho_0} + L_0 \rho_0 \cdot \frac{\bar{c} - c_0}{L_0 - \rho_0} \cdot \frac{1}{r}, \quad r \in (\rho_0, L_0),$$

$$v(t, r) = (1 - \gamma) \frac{\rho^2 \dot{\rho}}{r^2} = \frac{L^2 \dot{L}}{r^2}, \quad [Q - \gamma c(t, \rho)] \dot{\rho}(t) = D(T[t, \rho]) \left. \frac{\partial c}{\partial r} \right|_{\rho(t)},$$

$$D(T_L) \left. \frac{\partial c}{\partial r} \right|_{L(t)} = -b(T_L) c^2(t, L), \quad c(t, \rho(t)) = \bar{c} < Q, \quad T_L(t) \equiv T(t, L),$$

$$T(t, r) = A^\circ(t) + B^\circ(t) r^{-1}, \quad r \in [\rho(t), L(t)], \quad (1 - \gamma) [\rho_0^3 - \rho^3] = L_0^3 - L^3,$$

$$A^\circ(t) + B^\circ(t) L^{-1}(t) = T_L(t), \quad B^\circ(t) = \kappa \rho^2(t) \dot{\rho}(t), \quad T_L(t) = T_0 + vt.$$

Величины  $Q$ ,  $\bar{c}$ ,  $\kappa$ ,  $\gamma$  считаем постоянными в пределах ТДС-пика  $T \in [T^-, T^+]$ , коэффициенты диффузии и десорбции зависят от температуры по закону Аррениуса:  $D(T) = D_0 \exp\{-E_D/[RT]\}$ ,  $b(T) = b_0 \exp\{-E_b/[RT]\}$ .

После того как гидридное ядро исчезнет ( $\rho(t_*) = 0$ ), переходим к задаче дегазации без подвижных границ ( $r \in [0, L_*]$ ,  $L_* = L(t_*)$ )

с условием симметрии  $c_r(t, 0) = 0$  и равномерным прогревом металлических частиц  $T(t, r) = T_L(t) \equiv T(t)$ :

$$\frac{\partial c}{\partial t} = D(t) \left( \frac{\partial^2 c}{\partial r^2} + \frac{2}{r} \cdot \frac{\partial c}{\partial r} \right), \quad t \in (t_*, t^*), \quad r \in (0, L_*), \quad \varphi(r) = c(t_*, r),$$

$$D(t) \frac{\partial c}{\partial r} \Big|_{L_*} = -b(t)c^2(t, L_*), \quad \frac{\partial c}{\partial r} \Big|_{r=0} = 0, \quad D(t) \equiv D(T(t)), \quad b(t) \equiv b(T(t)).$$

Момент  $t^*$  определяется условием  $J(t^*) = b(t^*)c^2(t^*, L_*) \approx 0$ .

## 7. ПРЕОБРАЗОВАНИЕ КРАЕВОЙ ЗАДАЧИ

Введем замену переменных:  $(t, r) \leftrightarrow (t, x)$ ,  $r = \rho(t) + x[L(t) - \rho(t)]$ ,  $u(t, x) = c(t, r(t, x))$ ,  $w(t, x) = v(t, r(t, x))$ . Суть замены состоит в том, чтобы перевести подвижную область  $\rho(t) \leq r \leq L(t)$  в неподвижную  $0 \leq x \leq 1$ . Эта замена необходима для алгоритма численного решения, в противном случае возникают трудности с построением сетки по пространственной переменной.

Краевая задача ТДС-дегидрирования после замены:

$$\frac{\partial u}{\partial t} = \frac{1}{(L - \rho)^2} \frac{\partial}{\partial x} \left( D(T) \frac{\partial u}{\partial x} \right) + f(t, x) \cdot \frac{\partial u}{\partial x}, \quad t \in (0, t_*), \quad x \in (0, 1), \quad (1)$$

$$f(t, x) \equiv \left[ \dot{\rho}(1 - x) + \dot{L}x + \frac{2}{\rho + x(L - \rho)} \cdot D(T) - w(t, x) \right] \cdot \frac{1}{L - \rho},$$

$$w(t, x) \equiv v(t, r(t, x)) = (1 - \gamma) \cdot \frac{\rho^2 \dot{\rho}}{(\rho + x(L - \rho))^2} = \frac{L^2 \dot{L}}{(\rho + x(L - \rho))^2},$$

$$u(0, x) = \frac{L_0 u_0 - \bar{u} \rho_0}{L_0 - \rho_0} + L_0 \rho_0 \cdot \frac{\bar{u} - u_0}{L_0 - \rho_0} \cdot \frac{1}{\rho_0 + x(L_0 - \rho_0)},$$

$$u_0 \equiv u(0, 1), \quad \bar{u} = \bar{c}, \quad (1 - \gamma)[\rho_0^3 - \rho^3] = L_0^3 - L^3, \quad (2)$$

$$[Q - \gamma u(t, 0)] \dot{\rho}(t) = \left( \frac{D(T)}{L - \rho} \cdot \frac{\partial u}{\partial x} \right) \Big|_{x=0}, \quad \rho(0) = \rho_0, \quad u(t, 0) = \bar{u} < Q, \quad (3)$$

$$\left( \frac{D(T)}{L - \rho} \frac{\partial u}{\partial x} \right) \Big|_{x=1} = -b(T_L)u^2(t, 1), \quad T_L(t) \equiv T(t, L(t)) = T_0 + vt, \quad (4)$$

$$T(t, x) = A^\circ + B^\circ (\rho + x(L - \rho))^{-1}, \quad A^\circ + B^\circ L^{-1} = T_L, \quad B^\circ = \varkappa \rho^2 \dot{\rho}.$$

Формально можно забыть о физическом смысле  $\rho(t)$  как граници разделя фаз, а  $L(t)$  как радиуса частицы, и рассматривать их как функциональные параметры, подчиненные условиям (3) и (2). Получилась краевая задача для уравнения диффузии (1) (с коэффициентом, зависящим от параметров). Полученная краевая задача достаточно громоздка, но она в неподвижной области и ее можно решать, пользуясь обычными методами.

## 8. ПОСТРОЕНИЕ НЕЯВНОЙ РАЗНОСТНОЙ СХЕМЫ

Введем сетку  $x_i = ih$ ,  $i = 0, 1, \dots, N$ ,  $t_n = n\tau$ , где  $h = 1/N$  — шаг по радиусу,  $\tau$  — шаг по времени. Введем сеточные аппроксимации  $U_i^n \approx u(t_n, x_i)$ ,  $L_n \approx L(t_n)$ ,  $\rho_n \approx \rho(t_n)$ ,  $\dot{\rho}_n \approx \dot{\rho}(t_n)$ ,  $\dot{L}_n \approx \dot{L}(t_n)$ ,  $a_i^n \approx D(t_n, x_i - h/2)$ . Из начальных данных  $U_i^0 = u(0, x_i)$ . Для аппроксимации уравнения (1) с порядком  $O(h^2 + \tau)$  в узле  $(x_i, t_n)$  используем стандартный четырехточечный шаблон  $(x_{i\pm 1}, t_{n+1})$ ,  $(x_i, t_n)$ ,  $(x_i, t_{n+1})$ :

$$\begin{aligned} \frac{U_i^{n+1} - U_i^n}{\tau} &= \frac{a_{i+1}^{n+1} (U_{i+1}^{n+1} - U_i^{n+1}) - a_i^{n+1} (U_i^{n+1} - U_{i-1}^{n+1})}{h^2(L_{n+1} - \rho_{n+1})^2} + \\ &+ K_i^{n+1} \cdot \frac{U_{i+1}^{n+1} - U_{i-1}^{n+1}}{2h(L_{n+1} - \rho_{n+1})}, \\ K_i^{n+1} &\equiv \dot{\rho}_{n+1}(1 - x_i) + \dot{L}_{n+1} x_i + \frac{2a_i^{n+1}}{\rho_{n+1} + x_i(L_{n+1} - \rho_{n+1})} - w_i^{n+1}. \end{aligned}$$

Умножим обе части уравнения на  $(L_{n+1} - \rho_{n+1})^2 h^2$ , обозначим  $s_{n+1} = (L_{n+1} - \rho_{n+1})^2 h^2 / \tau$  и приведем подобные слагаемые:

$$\begin{aligned} b_i U_{i+1}^{n+1} - c_i U_i^{n+1} + a_i U_{i-1}^{n+1} &= -s_{n+1} U_i^n, \\ b_i &\equiv a_{i+1}^{n+1} + \frac{h}{2} K_i^{n+1} (L_{n+1} - \rho_{n+1}), \quad c_i \equiv a_i^{n+1} + a_{i+1}^{n+1} + s_{n+1}, \\ a_i &\equiv a_i^{n+1} - \frac{h}{2} K_i^{n+1} (L_{n+1} - \rho_{n+1}). \end{aligned}$$

Система решается методом прогонки:  $U_i^{n+1} = \alpha_{i+1}^{n+1} U_{i+1}^{n+1} + \beta_{i+1}^{n+1}$ ,  $i = 1, \dots, N - 1$ ,  $n \geq 1$ . Условия устойчивости выполнены ( $h \ll 1$ ):  $|c_i| \geq |a_i| + |b_i| = a_i^{n+1} + a_{i+1}^{n+1}$ ,  $s_{n+1} > 0$ . Границные условия имеют

вид

$$U_0^{n+1} = \eta Q, \quad D(T_L^{n+1}) \cdot \frac{3U_N^{n+1} - 4U_{N-1}^{n+1} + U_{N-2}^{n+1}}{2h(L_{n+1} - \rho_{n+1})} = -b(T_L^{n+1})(U_N^{n+1})^2.$$

Прогоночные коэффициенты для  $i = 2, \dots, N-1$  определяются формулами

$$\alpha_{i+1}^{n+1} = \frac{a_{i+1}^{n+1} + \frac{h}{2}K_i^{n+1}(L_{n+1} - \rho_{n+1})}{s_{n+1} + a_i^{n+1} + a_{i+1}^{n+1} - [a_i^{n+1} - \frac{h}{2}K_i^{n+1}(L_{n+1} - \rho_{n+1})]\alpha_i^{n+1}},$$

$$\beta_{i+1}^{n+1} = \frac{s_{n+1}U_i^n + [a_i^{n+1} - \frac{h}{2}K_i^{n+1}(L_{n+1} - \rho_{n+1})]\beta_i^{n+1}}{s_{n+1} + a_i^{n+1} + a_{i+1}^{n+1} - [a_i^{n+1} - \frac{h}{2}K_i^{n+1}(L_{n+1} - \rho_{n+1})]\alpha_i^{n+1}}.$$

Рассмотрим левую границу при  $i = 1$ . Подставляя в уравнение

$$s_{n+1}(U_1^{n+1} - U_1^n) = \left[ a_2^{n+1} + \frac{h}{2}K_1^{n+1}(L_{n+1} - \rho_{n+1}) \right] U_2^{n+1} - \\ - (a_1^{n+1} + a_2^{n+1})U_1^{n+1} + \left[ a_1^{n+1} - \frac{h}{2}K_1^{n+1}(L_{n+1} - \rho_{n+1}) \right] U_0^{n+1}$$

значение  $U_0^{n+1} = \eta Q$ , получим первые прогоночные коэффициенты

$$\alpha_2^{n+1} = \frac{a_2^{n+1} + \frac{h}{2}K_1^{n+1}(L_{n+1} - \rho_{n+1})}{a_1^{n+1} + a_2^{n+1} + s_{n+1}},$$

$$\beta_2^{n+1} = \frac{s_{n+1}U_1^n + [a_1^{n+1} - \frac{h}{2}K_1^{n+1}(L_{n+1} - \rho_{n+1})]\eta Q}{a_1^{n+1} + a_2^{n+1} + s_{n+1}}.$$

Вычисляем значения  $\alpha_2^{n+1}$  и  $\beta_2^{n+1}$ . Далее находим оставшиеся прогоночные коэффициенты  $\alpha_i^{n+1}$  и  $\beta_i^{n+1}$ ,  $i = 3, \dots, N$ . Для нахождения  $U_N^{n+1}$  используем граничное условие

$$D(T_L^{n+1}) \frac{3U_N^{n+1} - 4U_{N-1}^{n+1} + U_{N-2}^{n+1}}{2h(L_{n+1} - \rho_{n+1})} = -b(T_L^{n+1})(U_N^{n+1})^2.$$

Подставим связи  $U_{N-2}^{n+1} = \alpha_{N-1}^{n+1}U_{N-1}^{n+1} + \beta_{N-1}^{n+1}$ ,  $U_{N-1}^{n+1} = \alpha_N^{n+1}U_N^{n+1} + \beta_N^{n+1}$ . Положим  $A \equiv 3 - 4\alpha_N^{n+1} + \alpha_{N-1}^{n+1}\alpha_N^{n+1}$ ,  $B \equiv \alpha_{N-1}^{n+1}\beta_N^{n+1} + \beta_{N-1}^{n+1} - 4\beta_N^{n+1}$ . Граничное условие с учетом обозначений перепишется в форме

$$2h b(T_L^{n+1})(L_{n+1} - \rho_{n+1})D^{-1}(T_L^{n+1})(U_N^{n+1})^2 + AU_N^{n+1} + B = 0.$$

Это квадратное уравнение относительно  $U_N^{n+1}$ . Выбирается положительный корень. Значение функционального параметра  $\rho_{n+1} \approx \rho(t_{n+1})$  находим из разложения  $\rho_{n+1}$  в окрестности  $\rho_n$  по формуле Тейлора, учитывая только линейное слагаемое. Остальными слагаемыми пренебрегаем:

$$\rho_{n+1} \approx \rho_n + \tau \dot{\rho}_n \approx \rho_n + \frac{\tau D(T(t_n, \rho_n)) \{-3U_0^n + 4U_1^n - U_2^n\}}{2h(Q - \gamma \bar{u})(L_n - \rho_n)}.$$

Значение параметра  $L_{n+1}$  находим из соотношения  $(1 - \gamma)[\rho_0^3 - \rho^3] = L_0^3 - L^3$  при уже известном  $\rho_{n+1}$ .

**8.1. Диффузионная дегазация.** Гидридного ядра нет  $\rho(t_*) = 0$ . Это задача дегазации без подвижных границ ( $r \in [0, L_*]$ ,  $L_* = L(t_*)$ ) с условием симметрии  $c_r(t, 0) = 0$  и равномерным прогревом металлических частиц  $T(t, r) = T_L(t) \equiv T(t)$ :

$$\frac{\partial c}{\partial t} = D(t) \left( \frac{\partial^2 c}{\partial r^2} + \frac{2}{r} \cdot \frac{\partial c}{\partial r} \right), \quad t \in (t_*, t^*), \quad r \in (0, L_*), \quad \varphi(r) = c(t_*, r),$$

$$D(t) \frac{\partial c}{\partial r} \Big|_{L_*} = -b(t)c^2(t, L_*), \quad \frac{\partial c}{\partial r} \Big|_{r=0} = 0, \quad D(t) \equiv D(T(t)), \quad b(t) \equiv b(T(t)).$$

Коэффициенты для  $i = 1, \dots, N - 1$  упрощаются следующим образом

$$\alpha_{i+1}^{n+1} = \frac{1 + h/x_i}{s_{n+1} + 2 - \alpha_i^{n+1}[1 - h/x_i]}, \quad \beta_{i+1}^{n+1} = \frac{[1 - h/x_i]\beta_i^{n+1} + s_{n+1}U_i^n}{s_{n+1} + 2 - \alpha_i^{n+1}[1 - h/x_i]}.$$

Для того, чтобы найти первые прогоночные коэффициенты, пользуемся условием симметрии в центре шара:

$$\frac{\partial c}{\partial r}(t, 0) = 0 \Rightarrow \frac{2}{r}D(t)\frac{\partial c}{\partial r} = \frac{2}{r}D(t)\left(\frac{\partial c}{\partial r}(t, r) - \frac{\partial c}{\partial r}(t, 0)\right) \approx 2D(t)\frac{\partial^2 c(t, 0)}{\partial r^2}.$$

Таким образом уравнение диффузии при  $r \rightarrow +0$  перепишется в виде

$$\frac{\partial c}{\partial t} = 3D(t)\frac{\partial^2 c}{\partial r^2}.$$

В переменных  $(t, x)$  получаем уравнение ( $r = xL_*$ ,  $u(t, x) = c(t, r(x))$ ):

$$\frac{\partial u}{\partial t} = 3\frac{D(t)}{L_*^2} \cdot \frac{\partial^2 u}{\partial x^2}.$$

Разностная аппроксимация уравнения диффузии

$$\frac{U_i^{n+1} - U_i^n}{\tau} = 3 \frac{D(t_{n+1})}{L_*^2} \cdot \frac{U_{i+1}^{n+1} - 2U_i^{n+1} + U_{i-1}^{n+1}}{h^2}.$$

Умножим обе части уравнения на  $L_*^2 h^2$ ,  $s_{n+1} = h^2 L_*^2 (3\tau D(t_{n+1}))^{-1}$ , и приведем подобные слагаемые:

$$U_{i+1}^{n+1} - (2 + s_{n+1})U_i^{n+1} + U_{i-1}^{n+1} = -s_{n+1}U_i^n.$$

Границные условия имеют вид

$$\begin{aligned} \frac{\partial u}{\partial x} \Big|_{x=0} &= 0 \Rightarrow \frac{-3U_0^{n+1} + 4U_1^{n+1} - U_2^{n+1}}{2h} = 0, \\ D(t_{n+1}) \frac{3U_N^{n+1} - 4U_{N-1}^{n+1} + U_{N-2}^{n+1}}{2h L_*} &= -b(t_{n+1})(U_N^{n+1})^2. \end{aligned}$$

Рассмотрим левую границу при  $i = 1$ . Подставляя в уравнение

$$U_2^{n+1} - (2 + s_{n+1})U_1^{n+1} + U_0^{n+1} = -s_{n+1}U_1^n$$

соотношение  $(-3U_0^{n+1} + 4U_1^{n+1} - U_2^{n+1})/2h = 0$  или  $U_2^{n+1} = -3U_0^{n+1} + 4U_1^{n+1}$ , получаем первые прогоночные коэффициенты

$$\alpha_1^{n+1} = 1 - \frac{s_{n+1}}{2}, \quad \beta_1^{n+1} = \frac{s_{n+1}}{2} U_1^n.$$

Вычисляем значения  $\alpha_1^{n+1}$  и  $\beta_1^{n+1}$ . Далее находим оставшиеся прогоночные коэффициенты  $\alpha_i^{n+1}$  и  $\beta_i^{n+1}$ ,  $i = 2, \dots, N$ . Для нахождения  $U_N^{n+1}$  используем граничное условие

$$D(t_{n+1}) \frac{3U_N^{n+1} - 4U_{N-1}^{n+1} + U_{N-2}^{n+1}}{2h L_*} = -b(t_{n+1})(U_N^{n+1})^2.$$

Алгоритм определения  $U_N^{n+1}$  совпадает с приведенным выше.

## 9. АЛГОРИТМ ЧИСЛЕННОГО РЕШЕНИЯ

Рассматриваем краевую задачу ТДС-дегидрирования из п. 7 в переменных  $(t, x)$ .

1. Определяем начальные данные  $\psi(x) \equiv u(0, x)$ . Для этого находим значение концентрации  $u_0 = u(0, 1)$ , используя при  $t = 0$  граничное условие

$$\frac{D(T_L)}{L - \rho} \frac{\partial u}{\partial x} \Big|_{x=1} = -b(T_L)u^2(t, 1), \quad T_L(t) = T_0 + vt,$$

и начальное гиперболическое распределение ( $c(0, r) = A + B/r$ )

$$\psi(x) \equiv u(0, x) = \frac{L_0 u_0 - \bar{u} \rho_0}{L_0 - \rho_0} + \frac{L_0 \rho_0 (\bar{u} - u_0)}{(L_0 - \rho_0)[\rho_0 + x(L_0 - \rho_0)]}, \quad \bar{u} = \bar{c} < Q.$$

Получается квадратное уравнение с положительным корнем  $u_0$ :

$$u_0^2 + \frac{D(T_0)\rho_0}{b(T_0)L_0(L_0 - \rho_0)} u_0 - \frac{D(T_0)\rho_0 \bar{u}}{b(T_0)L_0(L_0 - \rho_0)} = 0.$$

2. Вычисляем начальную скорость движения границы  $\dot{\rho}_0 = \dot{\rho}(0)$  из условия Стефана

$$[Q - \gamma \bar{u}] \dot{\rho}_0 = \frac{D(T(0, 0))}{L_0 - \rho_0} \psi'(0), \quad \psi'(0) = -\frac{L_0(\bar{u} - u_0)}{\rho_0}.$$

Здесь коэффициент диффузии является нелинейной функцией от  $\dot{\rho}_0$ . Напомним, что температура  $T(t, r) = A^\circ(t) + B^\circ(t)/r$  в переменных  $(t, x)$  имеет выражение

$$T(t, x) = T_L(t) + \kappa \rho^2 \dot{\rho}(L - \rho)(1 - x)[L(\rho + x(L - \rho))]^{-1}, \quad T_L(t) = T_0 + vt.$$

Следовательно,  $T(0, 0) = T_0 + \kappa \dot{\rho}_0 \rho_0 (L_0 - \rho_0)/L_0 \equiv f_0(\dot{\rho}_0)$ . С учетом аррениусовской зависимости  $D(T)$  для  $\dot{\rho}_0$  получаем уравнение

$$\dot{\rho}_0 = -\frac{L_0(\bar{u} - u_0)}{(Q - \gamma \bar{u})(L_0 - \rho_0)\rho_0} D_0 \exp \left\{ -E_D/[Rf_0(\dot{\rho}_0)] \right\}.$$

Это задача о неподвижной точке ( $\dot{\rho}_0 = F_0(\dot{\rho}_0)$ ). В физически реальном диапазоне параметров модели (п. 10) она эффективно решается методом простой итерации.

3. Вычисляем  $\rho_1 = \rho_0 + \tau \dot{\rho}_0$  и значение  $L_1 \approx L(\tau)$  из  $(1 - \gamma)[\rho_0^3 - \rho_1^3] = L_0^3 - L_1^3$ . На отрезке  $r \in [\rho_1, L_1]$  определяем температуру  $T(\tau, r)$ , используя в выражении  $B^\circ(\tau) = \kappa \dot{\rho}_1 \rho_1^2$  вместо неизвестного пока  $\dot{\rho}_1$  начальное значение  $\dot{\rho}_0$ . В переменных  $(t, x)$  полагаем

$$T(\tau, x) = T_L(\tau) + \frac{\kappa \rho_1^2 \dot{\rho}_0 (L_1 - \rho_1)(1 - x)}{L_1[\rho_1 + x(L_1 - \rho_1)]}, \quad x \in [0, 1].$$

4. Неявным методом на слое  $t = \tau$  решаем диффузационную задачу.

5. Зная приближенное значение градиента концентрации  $\tilde{c}_r \approx c_r(\tau, \rho_1)$ , подставляем в уравнение движения границы (условие Стефана)  $[Q - \gamma \bar{c}] \dot{\rho}_1 = D(T(\tau, \rho_1)) \tilde{c}_r$  выражение температуры  $T(\tau, \rho_1)$  через переменную  $\dot{\rho}_1$  в силу соотношений

$$A^\circ(\tau) + B^\circ(\tau)/L(\tau) = T_L(\tau), \quad B^\circ(\tau) = \kappa \dot{\rho}_1 \rho_1^2.$$

В переменных  $(t, x)$  получаем уравнение для определения  $\dot{\rho}_1$ :

$$[Q - \gamma \bar{u}] \dot{\rho}_1 = \frac{D(f_1(\dot{\rho}_1))}{L_1 - \rho_1} \cdot \tilde{u}_x, \quad \tilde{u}_x \approx \left. \frac{\partial u}{\partial x} \right|_{x=0},$$

$f_1(\dot{\rho}_1) \equiv T_L(\tau) + \varkappa \rho_1 \dot{\rho}_1 (L_1 - \rho_1) L_1^{-1}$ . Задачу  $\dot{\rho}_1 = F_1(\dot{\rho}_1)$  решаем простыми итерациями,  $\dot{\rho}_0$  — начальное значение ( $\tau \ll 1$ ).

6. Возвращаемся к пункту 3, где мы вынуждены были использовать вместо  $\dot{\rho}_1$  значение  $\dot{\rho}_0$ . Теперь уже в  $B^\circ(\tau) = \varkappa \dot{\rho}_1 \rho_1^2$  подставляем текущее приближение  $\dot{\rho}_1$  и переходим к пункту 4. Таких итераций целесообразно сделать несколько. Для более грубых расчетов этот пункт алгоритма можно опустить.

7. Переходим на следующий временной слой  $t = 2\tau$  и действуем аналогично 3–6:  $\rho_2 = \rho_1 + \tau \dot{\rho}_1$  и т. д. Критерием остановки служит  $\rho(t_*) \approx 0$  ( $\rho < L_0/10$ ).

8. Переходим к этапу дегазации металла с растворенным  $H$ .

## 10. Вычислительные эксперименты

Будем ориентироваться на мощную вакуумную систему, пренебрегая возвратом на поверхность десорбированного водорода ( $\mu sr \approx 0$ ). На рисунках, иллюстрирующих влияние различных кинетических параметров, изображены графики плотности десорбции  $J(t) = b(T(t))c^2(t, L)$  при линейном нагреве. Значения параметров соответствуют убыванию максимума. Остальные параметры, общие для всех графиков, берутся из таблицы:

$b_0 = 7 \cdot 10^{-10} \text{ cm}^4/\text{s}$	$E_b = 145 \text{ KJ}$	$Q = 4 \cdot 10^{22} \text{ cm}^{-3}$	$\dot{T} = 0.3 \text{ K/s}$
$D_0 = 8 \cdot 10^{-3} \text{ cm}^2/\text{s}$	$E_D = 50 \text{ KJ}$	$L_0 = 5 \cdot 10^{-3} \text{ cm}$	$\eta = 0.15$
$T(0) = 400 \text{ K}$	$\gamma = 0.75$	$\bar{c} = \eta \cdot Q$	$\varkappa = 10^5$

Кружком отмечается момент окончания распада гидрида. Распределение частиц по радиусам нормальное: средний радиус частицы  $\bar{L} = 5 \cdot 10^{-3} \text{ см}$ ,  $\sigma = 10^{-3} \text{ см}$ . При численном моделировании использовался дискретный набор из 10 радиусов.

На рис. 1 приведен график зависимости  $J$  от текущей температуры  $T = T_L(t)$  ( $T_L(t) \leftrightarrow t$ ) и масштабированные кривые  $L \cdot \max J / [1.5 \max L]$ ,  $\rho \cdot \max J / [2 \max \rho]$ . На рис. 2 приведены графики  $J(T)$  ( $E_b = 150$ ). Изменения по убыванию максимума:  $L_0 = L_{\max} = 8 \cdot 10^{-3}$ ; с распределением частиц порошка по радиусам;  $L_0 = 5 \cdot 10^{-3}$ ;  $L_0 = L_{\min} = 2 \cdot 10^{-3}$ . Рис. 3 показывает, что  $E_b$

существенно влияет на динамику потока, меняется вид кривой. Вариации:  $E_b = 170; 150; 140$ . Рис. 4 отражает влияние энергии активации коэффициента распада гидрида  $E_k$  ( $k_0 = 5 \cdot 10^{-5}$ ). Вариации:  $E_k = 5; 10; 15$ . Чем медленнее диффузия (рис. 5, 6,  $\kappa = 10^6$ ), тем ниже и более смещен вправо график. Рис. 7 иллюстрирует влияние скорости нагрева. Рис. 8 отражает зависимость графика плотности десорбционного потока от коэффициента сжатия.

Таким образом, вариации параметров находятся в соответствии с физическими представлениями и модель может быть использована для идентификации кинетических параметров дегидрирования конкретных материалов.

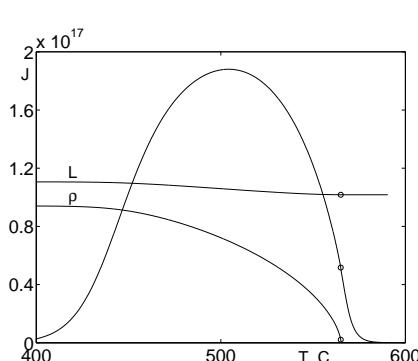


Рис. 1. Графики  $J$ ,  $L$ ,  $\rho$ .

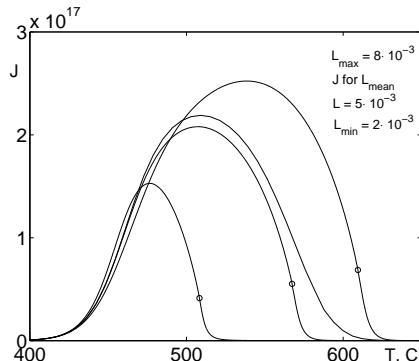


Рис. 2. Распределение по радиусам.

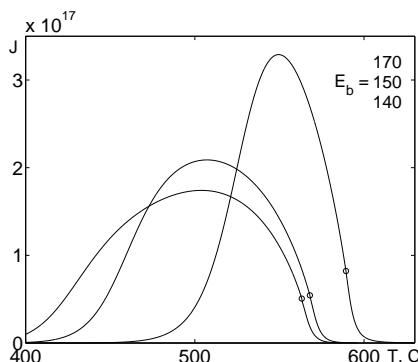


Рис. 3. Влияние параметра  $E_b$ .

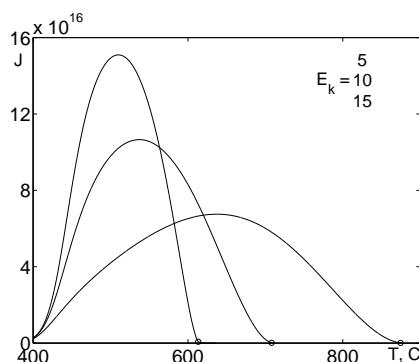


Рис. 4. Влияние параметра  $E_k$ .

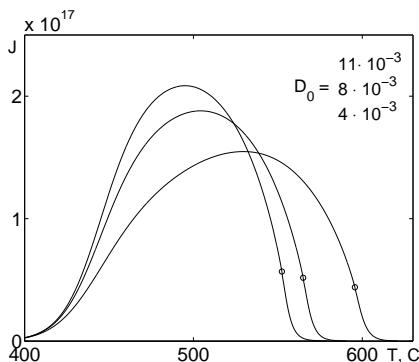
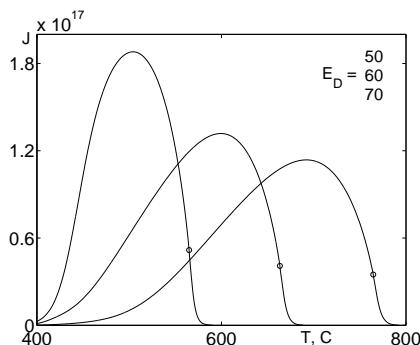
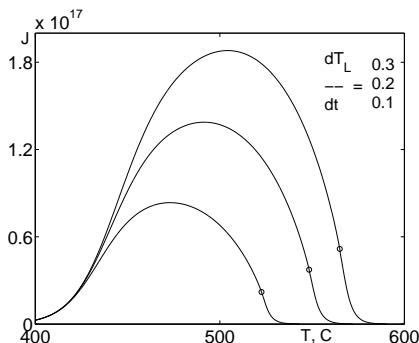
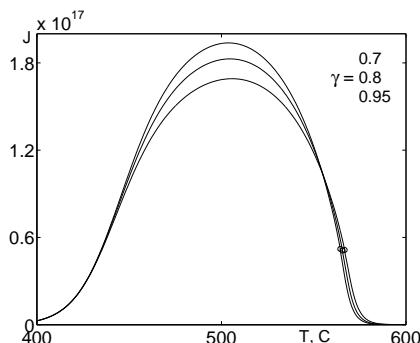
Рис. 5. Влияние параметра  $D_0$ .Рис. 6. Влияние параметра  $E_D$ .

Рис. 7. Влияние скорости нагрева.

Рис. 8. Влияние коэффициента  $\gamma$ .

## Литература

1. Заика Ю. В., Родченкова Н. И. Моделирование ТДС-спектра дегидрирования с учетом сжатия и теплопоглощения // Обозрение прикладной и промышленной математики, 2005, т. 12, вып. 4, с. 957–960.
2. Lufrano J., Sofronis P., Birnbaum H. K. Elastoplastically accommodated hydride formation and embrittlement// Journal of the Mechanics and Physics of Solids, 1998, v. 46, pp. 1497–1520.
3. Габис И. Е. Метод концентрационных импульсов для исследования транспорта водорода в твердых телах // Журнал технической физики, 1999, № 1, с. 99–103.
4. Castro F. J., Meyer G. Thermal desorption spectroscopy (TDS) method for hydrogen desorption characterization (I): theoretical aspects// Journal of Alloys and Compounds, 2002, v. 330–332, pp. 59–63.

5. Gabis I., Voit A., Evard E., Zaika Yu., Chernov I., Yartys V. Kinetics of Hydrogen Desorption from Powders of Hydrides of Metals // Journal of Alloys and Compounds, 2005, v. 404–406, pp. 312–316.
6. Zaika Yu. V., Chernov I. A. Nonlinear dynamical boundary-value problem of hydrogen thermal desorption // International Journal of Mathematics and Mathematical Sciences, 2003, № 23, pp. 1447–1463.
7. Заика Ю. В., Чернов И. А. Краевая задача с динамическими условиями и движущейся границей // Математическое моделирование, 2004, т. 16, № 4, с. 3–16.

УДК 519.837, 517.977

МОДЕЛИРОВАНИЕ ЭКОНОМИЧЕСКИХ ОТНОШЕНИЙ  
В ЛЕСНОМ КОМПЛЕКСЕ РЕСПУБЛИКИ КАРЕЛИЯ

А. Н. РЕТТИЕВА, А. В. РОДИОНОВ

Построена модель динамической игры управления лесными ресурсами. Получены оптимальные управление участников игры. Проведено численное моделирование для параметров, полученных по данным статистики по Республики Карелия.

1. ВВЕДЕНИЕ

Перед государством поставлена задача увеличения за 10 лет расчетной лесосеки в 2 раза. Каким образом эта проблема может быть решена? Мы предлагаем теоретико-игровой подход к решению, а именно исследуем динамическую игру двух лиц. В игре участвует центр (государство), которое регулирует уровень налогов и игрок (предприниматель), который вкладывает часть своей прибыли в развитие лесной отрасли.

2. ТЕОРЕТИКО-ИГРОВАЯ МОДЕЛЬ ОТРАСЛИ

Пусть  $x(t)$  – объем заготовок в момент времени  $t$ . Предприниматель имеет возможность производить три вида продукции: круглый лес, пиломатериалы и целлюлозно–бумажная продукция. Таким образом, он делит добываемое сырье на три части:  $k(t)x(t)$ ,  $l(t)x(t)$  и  $(1 - k(t)l(t))x(t)$ , где  $k(t)$ ,  $l(t)$  – его управления.

Прибыль, получаемая игроком в момент времени  $t$ , имеет вид

$$\begin{aligned} H(t) = & \{(c_1k(t)x(t) + c_2l(t)x(t) + c_3(1 - k(t) - l(t))x(t))(1 - \tau_0(t)) - \\ & - z_1k(t)x(t) - z_2l(t)x(t) - z_3(1 - k(t) - l(t))x(t)\}(1 - \tau(t)), \end{aligned}$$

где  $\tau(t)$  – налог на добавленную стоимость,  $\tau_0(t)$  – налог на прибыль,  $c_i$  – цены продажи единицы готовой продукции по видам,  $z_i$  – затраты на производство единицы продукции по видам ( $i = 1, 2, 3$ ).

Примем, что отношение объема заготовок на шаге  $t + \Delta t$  к объему на шаге  $t$ , равно отношению затрат, которые пошли на заготовку леса на шаге  $t + \Delta t$  к затратам на шаге  $t$ . При этом считаем, что предприниматель отдает  $\alpha(t)$  процентов своей прибыли на развитие производства, тогда

$$\frac{x(t + \Delta t)}{x(t)} = \frac{z_1 x(t) + \alpha(t)H(t)}{z_1 x(t)} \Delta t.$$

Откуда получим динамику развития объемов заготовок

$$x'(t) = \frac{\alpha(t)}{z_1} \{(c_1 k(t)x(t) + c_2 l(t)x(t) + c_3(1 - k(t) - l(t))x(t))(1 - \tau_0(t)) - z_1 k(t)x(t) - z_2 l(t)x(t) - z_3(1 - k(t) - l(t))x(t)\}(1 - \tau(t)), \quad x(0) = x_0.$$

Игрок заинтересован в максимизации своей рентабельности на промежутке  $[0, T]$ . Функционал выигрыша игрока имеет вид

$$I_1 = \int_0^T \frac{(1 - \alpha(t))H(t)}{z_1 k(t)x(t) + z_2 l(t)x(t) + z_3(1 - k(t) - l(t))x(t) + \alpha(t)H(t)}.$$

Прибыль, получаемая государством в момент времени  $t$ , имеет вид

$$G(t) = \tau_0(t)(c_1 k(t)x(t) + c_2 l(t)x(t) + c_3(1 - k(t) - l(t))x(t)) + \tau(t)\{(c_1 k(t)x(t) + c_2 l(t)x(t) + c_3(1 - k(t) - l(t))x(t))(1 - \tau_0(t)) - z_1 k(t)x(t) - z_2 l(t)x(t) - z_3(1 - k(t) - l(t))x(t)\} + \tau_l x(t) - M x(t),$$

где  $\tau_l$  – плата за лес,  $M$  – затраты на лесное хозяйство.

Государство параллельно с предпринимателем должно выделять деньги на развитие лесного хозяйства, пусть оно дает часть своей прибыли  $\beta(t)G(t)$ . Причем вклады двух игроков должны соответствовать друг другу, а именно должно выполняться соотношение

$$\frac{\alpha(t)H(t)}{z_1} = \frac{\beta(t)G(t)}{M}.$$

Откуда выражим  $\alpha(t) = \frac{\beta(t)G(t)z_1}{MH(t)}$ .

Примем, что государство также заинтересовано в максимизации своей рентабельности. Таким образом, получаем функционал выигрыша государства вида

$$I_2 = \int_0^T \frac{(1 - \beta(t))G(t)}{Mx(t) + \beta(t)G(t)} \Rightarrow \max .$$

С учетом соотношения для  $\alpha(t)$  перепишем функцию выигрыша предпринимателя как:

$$\begin{aligned} I_1 &= \int_0^T \frac{MH(t) - \beta(t)z_1G(t)}{M(z_1k(t)x(t) + z_2l(t)x(t) + z_3(1 - k(t) - l(t))x(t)) + \beta(t)z_1G(t)} \\ I_1 &\Rightarrow \max , \end{aligned}$$

и уравнение динамики имеет вид:

$$x'(t) = \frac{\alpha(t)}{z_1}H(t) = \frac{\beta(t)}{M}G(t), \quad x(0) = x_0 .$$

Примем, что налог на прибыль  $\tau_0$  фиксирован, тогда управления государства – это  $\tau(t)$  и  $\beta(t)$ , а управления предпринимателя –  $k(t)$  и  $l(t)$ .

Нас интересует оптимальное по Нэшу решение, а именно –  $(\tau^*(t), \beta^*(t), k^*(t), l^*(t))$ , такие что

$$\begin{cases} I_1(\tau^*, \beta^*, k^*, l^*) \geq I_1(\tau^*, \beta^*, k, l), \quad \forall k, l, \\ I_2(\tau^*, \beta^*, k^*, l^*) \geq I_2(\tau, \beta, k^*, l^*), \quad \forall \tau, \beta . \end{cases}$$

Для решения используем принцип максимума Понtryгина. Зафиксируем управление игрока (предпринимателя). Тогда Гамильтониан для центра (государства) имеет вид

$$H_2 = \frac{(1 - \beta(t))G(t)}{Mx(t) + \beta(t)G(t)} + \lambda_2 \frac{\beta}{M}G(t) .$$

Решая систему

$$\begin{cases} \frac{\partial H_2}{\partial \beta} = 0 \\ \frac{\partial H_2}{\partial \tau} = 0 \end{cases}$$

получим оптимальные управление следующего вида

$$\tau^*(t) = \frac{\tau_0 Sc + \tau_l - M}{-(1 - \tau_0)Sc + z_1k(t) + z_2l(t) + z_3(1 - k(t) - l(t))} ,$$

где  $Sc = c_1k(t) + c_2l(t) + c_3(1 - k(t) - l(t))$ ,

$$\beta^*(t) = \frac{1}{\lambda_2(t)x(t) - 1} .$$

Уравнение для  $\lambda_2$  получаем согласно принципу максимума

$$\lambda'_2 = -\frac{\partial H_2}{\partial x} = -\lambda_2 \frac{\beta}{M} \frac{\partial G}{\partial x}.$$

Центр заинтересован в увеличении выработки леса в  $k$  раз (у нас  $k = 2$ ), поэтому условие трансверсальности для сопряженной переменной имеет вид

$$\lambda_2(T) = b$$

и при выполнении  $b > 0$ , тождество  $b(kx(0) - x(T)) = 0$  выполняется при  $x(T) = kx(0)$  (закрепленный конец).

Подставляя найденное  $\tau^*(t)$  получим, что прибыль государства нулевая  $G(t) = 0$ . Таким образом, получим, что  $x'(t) = 0$ , т.е.  $x(t) = x_0$ , что противоречит поставленной задаче. Поэтому мы зафиксируем еще и  $\tau$  как параметр задачи. Тогда управление государства только  $-\beta(t)$ .

Снова решаем уравнение

$$\frac{\partial H_2}{\partial \beta} = 0$$

и получим оптимальные управления следующего вида

$$\begin{aligned} \beta_1^*(t) &= \frac{M}{(\tau\tau_0 - \tau_0 - \tau)Sc + \tau(z_1k + z_2l + z_3(1 - k - l)) - \pi + M} + \\ &+ \frac{\sqrt{(\lambda_2 x M ((\tau + \tau_0 - \tau\tau_0)Sc - \tau(z_1k + z_2l + z_3(1 - k - l)) + \pi)}}}{x\lambda_2((\tau\tau_0 - \tau_0 - \tau)Sc + \tau(z_1k + z_2l + z_3(1 - k - l)) - \pi + M)}, \\ \beta_2^*(t) &= \frac{M}{(\tau\tau_0 - \tau_0 - \tau)Sc + \tau(z_1k + z_2l + z_3(1 - k - l)) - \pi + M} - \\ &- \frac{\sqrt{(\lambda_2 x M ((\tau + \tau_0 - \tau\tau_0)Sc - \tau(z_1k + z_2l + z_3(1 - k - l)) + \pi)}}}{x\lambda_2((\tau\tau_0 - \tau_0 - \tau)Sc + \tau(z_1k + z_2l + z_3(1 - k - l)) - \pi + M)}. \end{aligned}$$

Зафиксируем управление центра. Гамильтониан игрока имеет вид

$$\begin{aligned} H_1 &= \frac{MH(t) - \beta(t)z_1G(t)}{M(z_1k(t)x(t) + z_2l(t)x(t) + z_3(1 - k(t) - l(t))x(t)) + \beta(t)z_1G(t)} + \\ &+ \lambda_1 \frac{\beta}{M} G(t). \end{aligned}$$

Решая систему

$$\begin{cases} \frac{\partial H_1}{\partial k} = 0 \\ \frac{\partial H_1}{\partial l} = 0 \end{cases}$$

получим две пары управлений  $-(k_1^*(t), l_1^*(t))$  и  $(k_2^*(t), l_2^*(t))$ . Приведем одну из них, т.к. выражения для второй пары слишком большие.

$$k_1^*(t) = \frac{(1-\tau)(1-\tau_0)z_1(c_3 - c_2)\beta(M - \tau_l) + \beta\tau(M - \tau_l)z_1(z_3 - z_2)}{(M(1-\tau_0)(1-\tau) - z_1\beta\tau)(c_1(z_2 - z_3) - c_2(z_1 - z_3) + c_3(z_1 - z_2))} + \\ + \frac{(c_3z_2 - z_3c_2)(\beta\tau z_1 - M(1-\tau)(1-\tau_0))}{(M(1-\tau_0)(1-\tau) - z_1\beta\tau)(c_1(z_2 - z_3) - c_2(z_1 - z_3) + c_3(z_1 - z_2))},$$

$$l_1^*(t) = \frac{(1-\tau)(1-\tau_0)z_1(c_1 - c_3)\beta(M - \tau_l) + \beta\tau(M - \tau_l)z_1(z_1 - z_3)}{(M(1-\tau_0)(1-\tau) - z_1\beta\tau)(c_1(z_2 - z_3) - c_2(z_1 - z_3) + c_3(z_1 - z_2))} + \\ + \frac{(c_1z_3 - z_1c_3)(\beta\tau z_1 - M(1-\tau)(1-\tau_0))}{(M(1-\tau_0)(1-\tau) - z_1\beta\tau)(c_1(z_2 - z_3) - c_2(z_1 - z_3) + c_3(z_1 - z_2))}.$$

Уравнение для  $\lambda_1$  получаем согласно принципу максимума

$$\lambda'_1 = -\frac{\partial H_1}{\partial x} = -\lambda_1 \frac{\beta}{M} \frac{\partial G}{\partial x}, \quad \lambda_1(T) = 0.$$

Далее подставляем найденные оптимальные управления в выражения для  $\beta_1^*(t)$  и  $\beta_2^*(t)$  и находим выражения для этих управлений, не содержащие  $k$  и  $l$ . Управлений будет два, обозначим их также  $\beta_1^*(t)$  и  $\beta_2^*(t)$ . Далее подставляем найденные стратегии в выражения для  $(k_1^*(t), l_1^*(t))$  и  $(k_2^*(t), l_2^*(t))$ . В итоге получим четыре вида управлений обоих игроков, а именно 1)  $\beta_1^*(t), (k_{11}^*(t), l_{11}^*(t))$ , 2)  $\beta_2^*(t), (k_{21}^*(t), l_{21}^*(t))$ , 3)  $\beta_1^*(t), (k_{12}^*(t), l_{12}^*(t))$ , 4)  $\beta_2^*(t), (k_{22}^*(t), l_{22}^*(t))$ . Полученные выражения сложны, поэтому здесь не приводятся.

### 3. ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

Моделирование проведено для следующих значений (полученных по данным статистики по Республике Карелия):

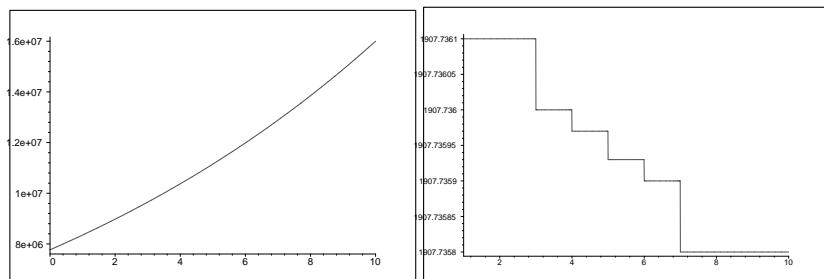
$$c_1 = 699.49, \quad c_2 = 2240.19, \quad c_3 = 3589.29, \\ z_1 = 756.14, \quad z_2 = 2280.52, \quad z_3 = 3269.84, \\ \tau_0 = 0.2, \quad \tau = 0.24, \quad \tau_l = 32.06, \\ M = 84.76, \quad x(0) = 7779900, \quad x(10) = 16000000.$$

Приведем пример для первой тройки управлений. Как видно из вида управлений, они не зависят от  $\lambda_1(t)$ . Поэтому решаем систему:

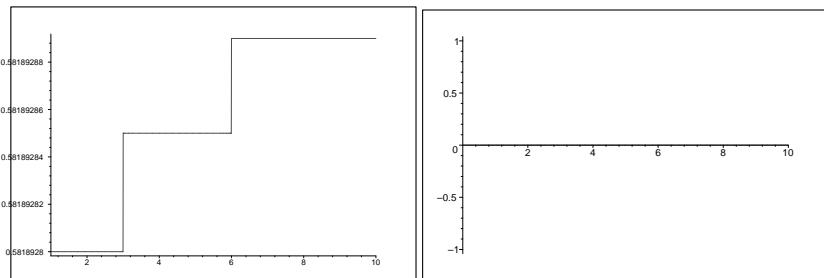
$$\begin{cases} x'(t) = \frac{\beta(t)}{M}G(t), & x(T) = 16000000, \\ \lambda'_2(t) = -\lambda_2(t) \frac{\beta(t)}{M} \frac{\partial G(t)}{\partial x}, & \lambda_2(T) = b. \end{cases}$$

При этом параметр  $b$  определяется из условия, что  $x(0) = 7779900$ .

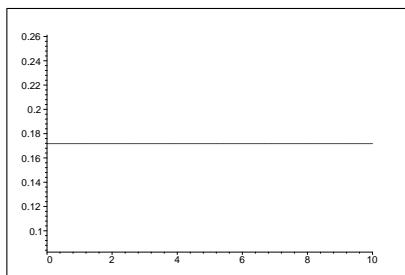
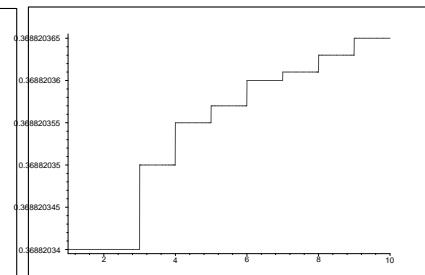
Оптимальное значение  $x^*(t)$  приведено на рис.1.1 (объем заготовок возрастает в 2 раза). На рис. 1.2 приведены значения средней цены за м3 сырья ( $Sc^*(t) = c_1 k^*(t) + c_2 l^*(t) + c_3(1 - k^*(t) - l^*(t))$ ).

Рис. 1.1. Значения  $x^*(t)$ Рис. 1.2. Значения  $Sc^*(t)$ 

На рис. 2.1 и 2.2 приведены оптимальные значения  $k$  и  $l$ . В данном примере мы получили, что предпринимателю необходимо 60% сырья использовать как круглый лес и 40% отдавать на целлюлозно-бумажную продукцию. При этом пиломатериалы не производятся вообще.

Рис. 2.1. Значения  $k^*(t)$ Рис. 2.2. Значения  $l^*(t)$

На рис. 3.1 и 3.2 приведены проценты прибылей государства и предпринимателя, которые необходимо направлять на развитие производства, для государства – это 17%, а для предпринимателя – 36%.

Рис. 3.1. Значения  $\beta^*(t)$ Рис. 3.2. Значения  $\alpha^*(t)$ 

## Литература

1. Basar T., Olsder G.J. Dynamic noncooperative game theory. Academic Press, New York, 1982.
2. Kaitala V., Hamalainen R.P., Ruusunen J. On the analysis of equilibria and bargaining in a fishery game // Optimal Control Theory and Economic Analysis, n.2, 1985, p. 593–605.
3. Mazalov V.V., Rettieva A.N. A fishery game model with age distributed population: reserved territory approach // Game Theory and Applications, v. 9, 2003, p. 56–72.
4. Mazalov V.V., Rettieva A.N. A fishery game model with migration: reserved territory approach // Game Theory and Applications, v. 10, 2004, p. 97–108.
5. Мазалов В.В., Реттиева А.Н. Об одной задаче управления популяцией // Обозрение прикладной и промышленной математики, № 9, вып.2, 2002, с. 293–306.
6. Мазалов В.В., Реттиева А.Н. Равновесие по Нэшу в задачах охраны окружающей среды // Математическое моделирование, том 18, № 5, 2006, с. 73–90.
7. Петросян А.А., Зенкевич Н.А., Семина Е.А. Теория игр. Москва, 1998.
8. Понtryagin L.S., Boltyanskiy V.G., Gamkrelidze R.V., Miščenko E.F. Matematicheskaya teoriya optimálnykh protsessov. M.: Nauka, 1976.
9. Родионов А.В. Повышение эффективности лесопользования. Петрозаводск, 2004.

УДК 512.544.43

**ПРИМЕНЕНИЕ ПАРАЛЛЕЛЬНЫХ ВЫЧИСЛЕНИЙ ДЛЯ  
РЕШЕНИЯ НЕКОТОРЫХ ЗАДАЧ МЕХАНИКИ ТВЕРДОГО  
ДЕФОРМИРУЕМОГО ТЕЛА**

М.П. ЩЕРБАК

В работе представлен пример применения параллельных вычислений для решения одной из задач механики твердого деформируемого тела. Речь идет о задаче определения усилий и перемещений в механической системе, состоящей из набора абсолютно жестких тел, жестких стержней и податливых односторонних связей.

Определяющие соотношения задачи приводят к задаче квадратичного программирования, решать которую предлагается методом Вулфа. Все вычисления производятся на клавестре — наиболее доступном суперкомпьютере.

В настоящей работе рассматривается расчет механической системы абсолютно жестких тел, жестких стержней и податливых односторонних связей на параллельном компьютере. Рассматриваемая задача находится на стыке двух актуальных направлений разных областей знаний: механики твердого деформируемого тела и информационных технологий. Для первой актуальным направлением является расчет систем с односторонними связями, который в последнее время вызывает все больший научный и практический интерес. Параллельные вычисления, с другой стороны, являются сегодня одним из приоритетных направлений исследований области информационных технологий в связи с началом массового производства многоядерных процессоров.

Односторонней связью называют такой объект, который может воспринимать усилия только одного знака. Примером такой

связи может служить обычная нить: она сопротивляется растяжению, но не препятствует сближению своих концов. Про такую связь говорят, что она работает только на растяжение.

Учет односторонних связей приводит к появлению неравенств в математической модели механической системы. Если принять правило знаков для перемещения по направлению односторонней связи  $Z$  и усилия в ней  $X$  таким, что  $X \geq 0$ ,  $Z \geq 0$ , то условия работы в связи можно записать в виде:

$$X \geq 0, \quad Z \geq 0, \quad XZ = 0. \quad (1)$$

Применение одного из известных методов строительной механики (метода сил) к рассматриваемой механической системе дает следующие соотношения:

$$\begin{aligned} N &= sX + S, \\ Z &= aX - s^\top \Delta + A, \\ X &\geq 0, \quad Z \geq 0, \\ N &\geq 0, \quad \Delta \geq 0, \end{aligned} \quad (2)$$

$$X \cdot Z = 0 \quad N \cdot \Delta = 0. \quad (3)$$

Здесь векторы  $X$ ,  $N$  имеют смысл усилий в связях рассматриваемой механической системы, а векторах  $Z$  и  $\Delta$  записаны перемещения, сопряженные по работе с  $X$  и  $N$  соответственно.

Уравнения (2) и (3) являются условиями Куна-Таккера (необходимыми и достаточными условиями существования решения) для следующей квадратичной задачи:

$$\min\left\{\frac{1}{2} X a X + AX \mid N = sX + S, \quad N \geq 0, \quad X \geq 0\right\}. \quad (4)$$

Матрица  $a$  в программе (4) является симметричной положительно определенной, что позволяет применять для решения этой задачи любой метод квадратичного программирования. Предпочтение, однако, было отдано методу Вулфа. Во первых, — это точный метод, который позволяет прийти к цели за конечное число шагов. Во-вторых, он использует удобную вычислительную схему симплексного метода.

Метод Вулфа состоит из трех этапов, на каждом из которых выполняется примерно одинаковый объем вычислений. На первых двух этапах осуществляется поиск базисного решения, а на последнем — собственно решения задачи. При этом везде применяется алгоритм симплекс метода с учетом дополнительного правила выбора ведущего элемента, которое говорит о том, что обе величины пары  $X_i Z_i$  не могут быть базисными одновременно.

Если для решаемой системы можно было бы указать базисное решение заранее, то это позволило бы сократить требуемый объем вычислений в несколько раз. В работе [1] показана такая возможность для рассматриваемого случая. Для этого осуществляется переход к двойственной задаче: в качестве функции цели выбирается параметр  $p$  внешней нагрузки на конструкцию. Если в исходной задаче требовалось отыскать минимум энергетической функции при косвенных ограничениях на параметр  $p$  ( $p \leq 1$ ), то теперь ставится задача на максимум  $p$  при косвенных ограничениях на энергетических функциях:

$$\max\{p \mid \eta = p^* - p \geq 0; \quad p \geq 0\}. \quad (5)$$

В статье [2] получены условия Куна-Таккера для этой программы:

$$\begin{aligned} N &= sX + pS^* + S^0, \\ Z &= aX - s^\top \Delta + pA^* + A^0, \\ X &\geq 0, \quad Z \geq 0, \\ N &\geq 0, \quad \Delta \geq 0, \\ X \cdot Z &= 0 \quad N \cdot \Delta = 0. \end{aligned} \quad (6)$$

Векторы  $A^0$  и  $S^0$ , содержащие величины начальных усилий и перемещений в односторонних связях, благодаря выбранному правилу знаков, всегда неотрицательны. Таким образом, симплекс-таблица, записанная для программы (5), всегда реализует базисное решение: при нулевых значениях переменных  $p$ ,  $X$  ни одно из условий Куна-Таккера (6) не оказывается нарушенным.

Размер симплексной таблицы, над которой производятся вычисления —  $(2 + n + l) \times (2 + n + l)$ , где  $n$  и  $l$  — длины векторов  $Z$

и  $N$  соответственно. При этом величина  $m = n + l$  равна числу односторонних связей в системе.

Алгоритм не позволяет определить заранее требуемый объем вычислений, т. к. это зависит не только от размера, но еще и от содержания симплексной таблицы. Однако практика расчетов механических задач показала, что обычно для нахождения решения требуется выполнить около  $m$  шагов (операций исключения Гаусса-Жордана). Из этого следует оценка сложности всего алгоритма  $O(m^3)$ . При такой трудоемкости временные затраты на вычисления заметны уже при  $m = 1000$  даже при использовании современных персональных компьютеров. Так, на компьютере с процессором Intel Celeron 2100 поиск решения задачи с  $m = 1000$  занял 8 секунд, а задачи с  $m = 2000$  — уже 55. При этом на алгоритм Вулфа приходилось более 95% от общего объема вычислений.

Подобные показатели неприемлемы в случаях, когда требуется выполнить серию расчетов, например, при анализе движущейся системы. Особенно требовательными к скорости вычислений являются приложения, выполняющие симуляцию движения механической системы в режиме реального времени. Для таких программ необходимо предусмотреть способы увеличения производительности, причем степень увеличения должна быть, по возможности, конфигурируемой.

Подобную задачу можно решить прибегнув к параллельным вычислениям, реализуемых на масштабируемых суперкомпьютерах. Было решено разрабатывать алгоритм для кластеров — наиболее доступного типа суперкомпьютеров, который относится к классу параллельных компьютеров с распределенной памятью. Предлагаемый алгоритм учитывает главные особенности программ для подобных компьютеров, такие как:

- данные задачи распределяются по вычислительным узлам;
- каждый узел производит вычисления только с локальными данными;
- обмен данными между узлами осуществляется через локальную вычислительную сеть посредством программного интерфейса MPI.

Первое означает, что все данные задачи по возможности равномерно распределяются между узлами кластера. Каждый из них получает только часть задачи — такую, чтобы с ней можно было выполнить максимально возможный объем вычислений без обращения к данным других узлов. Это требование связано с тем, что обращение к удаленным данным напрямую невозможно, так как вычислительный узел кластера не имеет прямого доступа к локальной памяти другого узла. Обмен данными осуществляется косвенно, посредством посылки и приема сообщений через локальную сеть.

В обычном алгоритме Вулфа большая часть вычислительной мощности приходится на многократную операцию исключения Гаусса-Жордана над симплексной таблицей. Оставшуюся часть занимает процедура выбора ведущих элементов. В статье [3] был получен параллельный алгоритм, который выполняет две эти процедуры и отвечает требованиям, приведенным выше.

Для выявления параллелизма в существующем алгоритме Вулфа была применена одна из методик, предложенных В. В. Воловодиным [4]. Информационный граф (граф, показывающий последовательность информационных связей) одного итерационного шага алгоритма совмещался с разными вариантами схем распределенного хранения данных. Среди множества вариантов был выбран способ хранения по блокам-столбцам. На рис. 1 показана выбранная схема.

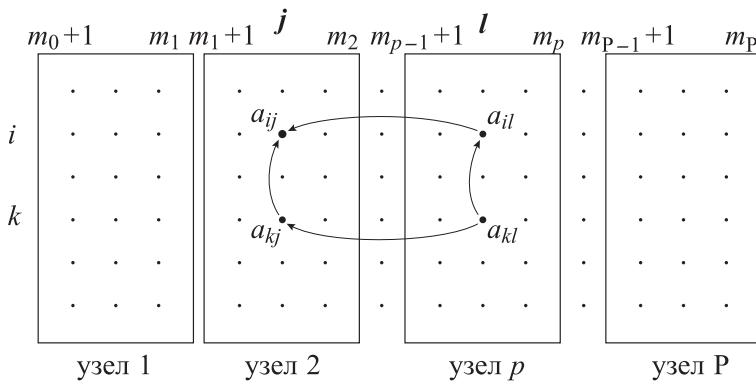


Рис. 1

В индексных обозначениях алгоритм может быть записан сле-

дующим образом:

$$\begin{aligned}
 a_{kj} &= -a_{kj}/a_{kl}, & j &= m_{p-1}, m_{p-1} + 1, \dots, l - 1, l + 1, \dots, m_p, \\
 a_{il} &= a_{il}/a_{kl}, & i &= 1, 2, \dots, k - 1, k + 1, \dots, n \\
 a_{ij} &= a_{ij} + a_{il} \cdot a_{kj}, & i &= 1, 2, \dots, k - 1, k + 1, \dots, n, \\
 && j &= m_{p-1}, m_{p-1} + 1, \dots, l - 1, l + 1, \dots, m_p, \\
 a_{kl} &= 1/a_{kl}.
 \end{aligned} \tag{8}$$

Если  $l \geq m_{p-1}$  или  $l \leq m_p$ , то

$$\begin{aligned}
 a_{kj} &= -a_{kj}/a_{kl}, & j &= m_{p-1}, m_{p-1} + 1, \dots, m_p, \\
 a_{il} &= a_{il}/a_{kl}, & i &= 1, 2, \dots, k - 1, k + 1, \dots, n \\
 a_{ij} &= a_{ij} + a_{il} \cdot a_{kj}, & i &= 1, 2, \dots, k - 1, k + 1, \dots, n, \\
 && j &= m_{p-1}, m_{p-1} + 1, \dots, m_p, \\
 a_{kl} &= 1/a_{kl}.
 \end{aligned} \tag{9}$$

Из рис. 1 и из приведенного алгоритма видно, что на каждом шаге метода Вулфа на любом вычислительном узле используется только один нелокальный элемент — столбец  $a_{il}$  (этот элемент локален только для одного узла  $p$ , для которого  $m_{p-1} \leq l \leq m_p$ ). Так как данные, требуемые всеми узлами, одинаковы, они передаются за один раз, широковещательно перед началом вычислительного шага.

Процедура параллельного поиска ведущего элемента построена на тех же рассуждениях, что и предыдущая операция. Главное отличие заключается в том, что здесь роль столбца  $a_{il}$  выполняет столбец  $a_{il}$ , т. е. столбец с нелокальными данными известен заранее. Это позволяет рассыпать столбец  $a_{il}$  однократно до начала вычислений. В процессе вычислений на всех узлах над ним выполняются одинаковые операции, в результате чего любой процесс постоянно имеет локальный доступ к текущим значениям вектора свободных членов. Таким образом удается избежать межузловых коммуникаций, связанных с пересылкой массива  $a_{il}$ .

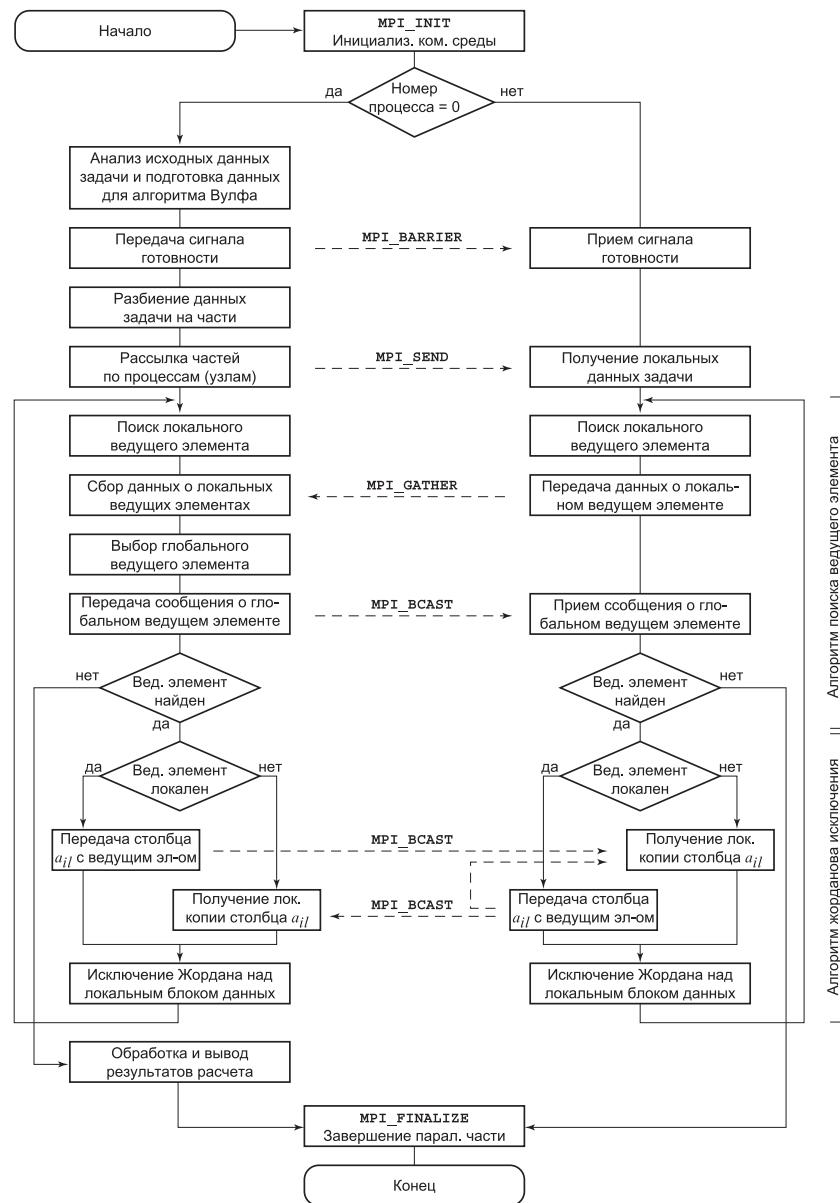


Рис. 2

На рис. 2 представлена блок-схема параллельной части программы. Она начинается с вызова функции MPI\_INIT, которая инициализирует коммуникационную среду. После ее вызова каждый процесс кластера получает уникальный номер-адрес, который в терминологии MPI называется номером процесса. Сразу после инициализации коммуникационной среды происходит ветвление программы: программа главного процесса (узла) с номером 0 движется по левой ветви алгоритма (см. рис. 2), а программы остальных – по правой. Передача сообщений происходит в основном между правой и левой ветвями алгоритма. На блок-схеме межузловой обмен данными показан пунктирными линиями. Над каждой из них приведено имя MPI-функции, с помощью которой осуществляется этот обмен.

Описанный выше алгоритм был реализован в программе расчета систем абсолютно жестких тел и податливых односторонних связей. В статье [3] описан численный эксперимент с этой реализацией, по результатам которого были построены графики производительности и масштабируемости вычислительной системы.

Вычисления производились на кластере, созданном на базе одного из компьютерных классов Петрозаводского университета. Кластер состоит из 6 компьютеров, которые имеют следующую конфигурацию: процессор – Intel Celeron 2100, оперативная память – 256 Mb (DDR), сетевой адаптер – Fast Ethernet 100 Mbit, ОС – Linux Mandrake 10.0, библиотеки MPI – MPICH 1.2.5, топология сети – звезда.



Рис. 3

В качестве тестовой задачи была взята задача строительной механики, в которой требуется определить усилия в односторонних связях механической системы. Она состоит из двух абсолютно жестких тел, которые соединены между собой односторонними податливыми связями, работающими только на растяжение. Связи устанавливаются случайным образом, а количество их произвольно. Всего было сгенерировано 9 расчетных схем с разным числом связей. Все тестовые системы были рассчитаны в четырех конфигурациях кластера: 1 работающий узел, 2, 4 и 6 узлов.

В каждом эксперименте измерялось время, затраченное на вычисления. По этим данным были построены графики производительности и масштабируемости.

Производительность одноузловой установки, достигнув насыщения при  $n > 1000$ , колебалась около отметки 430 MFLOPS. При этом теоретическая пиковая производительность процессора, установленного во всех узлах (Intel Celeron 2100), равна 2100 MFLOPS. Реальная скорость вычислений, таким образом, на одном процессоре составила всего лишь 20,4% от теоретической. Подобный показатель можно считать нормальным для неоптимизированного кода, оперирующего с большими массивами вещественных чисел. Для сравнения, в teste LAPACK, собранном с неоптимизированной reference-версией библиотеки низкоуровневых функций BLAS, реальная производительность одного процессора в некоторых расчетах ниже пиковой в 3 раза. С помощью оптимизации, учитывающей архитектуру всего вычислительного узла, строение его процессора, размеры кэшей и т. п., реальную скорость удается довести до 80 – 100% от теоретической.



Рис. 4

Вторым после производительности важным показателем эффективности кластерных систем является масштабируемость. Эта характеристика показывает способность кластера наращивать свою производительность в одной и той же задаче с подключением к вычислениям новых узлов. Обычно масштабируемость представляют в виде графика зависимости производительности от размера кластера (количество узлов). Рядом с этой кривой в тех же осях строят график суммы производительностей узлов кластера, который считается линией идеальной масштабируемости.

Результаты проведенных тестов показывают преимущества использования параллельных вычислений и говорят о том, что предложенный параллельный алгоритм может быть эффективно применен для решения больших задач. Его реализация показала хорошую масштабируемость на оборудовании, не предназначенном для построения кластеров, что указывает на высокие «параллельные» качества алгоритма. Применение современной аппаратной базы позволяет использовать эти качества в полной мере и сделает возможным применение предложенного алгоритма для решения задач с очень большой размерностью.

## Литература

1. Перельмутер А. В. Использование методов квадратичного программирования для систем с односторонними связями // Исследования по теории сооружений. Вып. XIX. М.: Стройиздат, 1972, с. 138–147.
2. Гольдштейн Ю. Б., Щербак М. П. Разрешающие уравнения для системы жестких тел и податливых связей // Рукопись деп. в ВИНИТИ 02.04.2003, 587-B2003.
3. Щербак М. П. Параллельный алгоритм метода Вулфа для суперкомпьютеров кластерной архитектуры // Рукопись деп. в ВИНИТИ 17.12.2005, 1609-B2005.
4. Воеводин В. В., Воеводин Вл. В. Параллельные вычисления. СПб.: БХВ-Петербург, 2002.

## ABSTRACTS

### **Section I. METHODS OF MATHEMATICAL MODELLING**

*Abakumov Y.G., Banin M.V.* On sets of convergences guaranteed by test type theorems.

The sets in the space  $C[a, b]$  on which the sequence of some linear operator classes is convergent on the test set implies that the sequence is convergent on the whole set. It is stated that these sets do not increase with the preservation of convergences. The general scheme for the sequence of linear functionals in the Banach space is considered.

*Ivanov A.P., Ostov Yu.Ya.* Duality and principle of expansion in the problem of dynamics of flight.

The paper considers the problem of maximizing kinetic energy in the end point of the path of a flying object. The solution method is construction of sub-optimal control.

*Popov V.V.* Modelling of heat-mass transfer processes in hydride accumulator of hydrogen.

The one-dimensional model of heat-mass transfer processes in hydride accumulator of hydrogen is considered. The model coefficients depend on temperature and concentration. The chemical reactions of hydrogenation and dehydrogenation are taken into account.

*Popov V.V., Semin P.A.* Implementation and use of modelling problems of hydrogen interaction with solids in special software.

The practical aspects of implementation and use of various models of hydrogen interaction with solids in special software are considered.

*Stafeev S.V.* On the parameter identifiability of a factor analysis model with correlated residuals and correlated factors.

A factor analysis model with correlated residuals and correlated factors is considered. Conditions of local parameter identifiability are obtained and the class of models satisfying these conditions is found.

*Tararin V.M., Cheplyukova I.A.* On the action of a random s-mapping of a finite Boolean lattice.

s-mapping of a finite Boolean lattice  $L$  is the upper endomorphism  $\varphi$  of lattice  $L_n$  such that from  $h(x) \leq 1$  follows  $h(\varphi(x)) = h(x)$ . In the paper we consider the action of a random s-mapping of a finite Boolean lattice with a length  $n$  as  $n \rightarrow \infty$ .

*Falko A.A.* A best-choice game with the possibility of an applicant refusing an offer and with redistribution of probabilities.

A full-information best-choice game is considered. The optimal strategies in case of the possibility of an applicant refusing an offer and the redistribution of probabilities are obtained.

*Khvorostyanskaya E.V.* On emergence of the giant  $\varphi$ -primitive ideal of a random automorphism  $\varphi$  of a finite Boolean lattice.

We consider the set  $A_{n,N}$  of all automorphisms  $\varphi$  of a Boolean lattice with a length  $n$  such that the number of all  $\varphi$ -primitive ideals is equal to  $N$ . We specify the uniform distribution on  $A_{n,N}$ . It is proved that the giant  $\varphi$ -primitive ideal of a random automorphism  $\varphi \in A_{n,N}$  can occur only if  $n, N \rightarrow \infty$  so that  $N/\ln n \rightarrow 0$ .

*Chernov I.A.* Convergence of the lattice approximations of the solution to the boundary-value problem with nonlinear boundary condition and moving bound.

In the paper we consider the Stefan-type boundary-value problem with moving bound and nonlinear boundary condition. It appears as a model of hydrating under constant conditions. We present the numerical method; it provides lattice approximations of  $\sqrt{h}$  precision,  $h$  is for the lattice step. We prove that these approximations converge to a weak continuous solution to the PDE; this solution also satisfies the boundary conditions. The moving bound belongs to  $C_1$ . So we can also claim the existence of a weak solution to the considered problem.

## Section II. COMPUTER SCIENCE

*Borodina A.V., Morozov E.V.* Overflow probability estimation with accelerated regeneration simulation in the queue  $M/M/1$ .

This paper is devoted to a modification of speed-up method Splitting. This approach allows to use the regenerative simulation techniques to construct the confidence interval for the overflow probability.

*Kiseleva N.S.* On an "Internet" type, random graph model.

Random graphs, where vertex degrees are independent random variables are considered. The distribution of the vertex degrees follows the  $\mathbf{P}\{D = d\} = d^{-\tau}/\zeta(\tau)$  law, where  $D$  is the random variable denoting the vertex degree,  $\zeta(\tau) = \sum_{d=1}^{\infty} d^{-\tau}$  is Riemann zeta function,  $d = 1, 2, \dots$ ,  $\tau \in (2, 3)$  (finite mean and infinite variance). Such graphs can be used as models for Internet topology. Limit theorems for the extreme terms of the variation series formed from the nondecreasing sequence of the vertex degrees and for the sum of the node degrees when  $N \rightarrow \infty$  are proved.

*Kukin V.D.* A composite evolutionary algorithm for solving the weighted Steiner tree problem.

A composite evolutionary algorithm for solving the weighted Steiner tree problem is suggested. Main results of testing the algorithm are reported.

*Lazutina A.A.* On the problem of optimal control of three parallel stacks in bounded space.

The paper considers issues related to building mathematical models and optimal algorithms of consecutive and connected three parallel stacks control in a single-level memory. These models were constructed as random walks.

*Pechnikov A.A.* The problem of rational allocation of links in a controlled and localized Internet-resource system.

The problem of rational allocation of cross-links to Internet-resource within a some localized system is considered. Each participant is obliged to make a certain number of links to other participants to increase their importance so that the final values of the importance differed from the value averaged for all participants of the system as little as possible (a case of a "social" decision).

*Rodchenkova N.I.* A differences scheme for modelling the TDS-spectrum of dehydrating of metals with heat absorption and a size reduction effect.

In the paper we consider the problem with moving bounds that models the kinetics of hydrogen desorption from hydrides of metals. Phase change, size reduction, desorption processes on the surface, and heat absorption on the phase bound are taken into consideration. The differences scheme for the corresponding boundary-value problem

with moving bounds is presented for the experimental method of thermal desorption spectrometry (TDS). The results of numerical modeling are presented.

*Rettieva A.N., Rodionov A.V.* The model of dynamic game of forest resources management problem is constructed.

The equilibrium controls of the participants are received. The numerical modeling for parameters is carried out according to statistics on Karelia.

*Shcherbak M.P.* Parallel computing in application to analysis of some problems in mechanics.

One problem of mechanics is solved in this paper by means of parallel computing. More precisely, in this paper one can find a description of the analysis of a mechanical system composed of rigid bodies and unilateral constrains. This analysis is performed with the help of parallel computing.

Defining relations of this problem related to the quadratic programming problem, that is solved by Vulf method. Parallel calculations are executed on the cluster which is the most available type of scalable parallel computer.

Научное издание

**МЕТОДЫ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ  
И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

**Выпуск 7**

*Печатается по решению Ученого совета  
института прикладных математических исследований  
Карельского научного центра РАН*

Оригинал-макет подготовлен Е.Н. Спектор

Изд. лиц. № 00041 от 30.08.99. Подписано в печать 05.10.06.  
Формат 60x84 1/16. Бумага офсетная. Гарнитура «Times». Печать офсетная.  
Уч.-изд. л. 13,5. Усл. печ. л. 12,8. Тираж 150 экз. Изд. № 66. Заказ 607.

Карельский научный центр РАН  
185003, Петрозаводск, пр. А. Невского, 50  
Редакционно-издательский отдел