

ISSN 1819-3161

ПРОБЛЕМЫ УПРАВЛЕНИЯ

3/2018

CONTROL  SCIENCES

Редакционный совет

Акад. РАН С.Н. Васильев, акад. РАН С.В. Емельянов, акад. РАН И.А. Каляев, акад. РАН В.А. Левин, чл.-корр. РАН Н.А. Махутов, акад. РАН Е.А. Микрин, чл.-корр. РАН П.П. Пархоменко, чл.-корр. РАН А.Ф. Резчиков, акад. РАН Е.А. Федосов

Редколлегия

Д-ра техн. наук Ф.Т. Алескеров, В.Н. Афанасьев, Н.Н. Бахтадзе, канд. техн. наук Л.П. Боровских (зам. гл. редактора), д-ра техн. наук В.Н. Бурков, В.М. Вишневский, д-р экон. наук М.И. Гераськин, д-р техн. наук А.А. Дорофеюк, д-р экон. наук В.В. Клочков, д-ра техн. наук С.А. Краснова, О.П. Кузнецов, В.В. Кульба, А.П. Курдюков, д-ра физ.-мат. наук А.Г. Кушнер, А.А. Лазарев, д-р техн. наук В.Г. Лебедев, д-р психол. наук В.Е. Лепский, д-р техн. наук А.С. Мандель, д-р биол. наук А.И. Михальский, чл.-корр. РАН Д.А. Новиков (гл. редактор), д-ра техн. наук Б.В. Павлов, Ф.Ф. Пашенко (зам. гл. редактора), д-р физ.-мат. наук Л.Б. Рапопорт, д-р экон. наук С.В. Ратнер, д-ра техн. наук Е.Я. Рубинович, В.Ю. Рутковский, д-р физ.-мат. наук М.В. Хлебников, д-р техн. наук А.Д. Цвиркун, д-р физ.-мат. наук П.Ю. Чеботарёв, д-р техн. наук И.Б. Ядыкин

Руководители региональных редакционных советов

Владивосток – д-р техн. наук О.В. Абрамов (ИАПУ ДВО РАН)
Волгоград – д-р физ.-мат. наук А.А. Воронин (ВГУ)
Воронеж – д-р техн. наук С.А. Баркалов (ВГАСУ)
Курск – д-р техн. наук С.Г. Емельянов (ЮЗГУ)
Липецк – д-р техн. наук А.К. Погодаев (ЛГТУ)
Пермь – д-р техн. наук В.Ю. Столбов (ПНИПУ)
Ростов-на-Дону – д-р техн. наук Г.А. Угольницкий (ЮФУ)
Самара – д-р техн. наук В.Г. Засканов (СГАУ)
Саратов – д-р техн. наук В.А. Твердохлебов (ИПТМУ РАН)
Уфа – д-р техн. наук Б.Г. Ильясов (УГАТУ)



CONTROL SCIENCES

**Научно-технический
журнал**

6 номеров в год

ISSN 1819-3161

Издается с 2003 года

УЧРЕДИТЕЛЬ

Федеральное государственное
бюджетное учреждение науки
Институт проблем управления
им. В.А. Трапезникова РАН

Главный редактор
чл.-корр. РАН

Д.А. Новиков

Заместители главного
редактора

Л.П. Боровских, Ф.Ф. Пащенко

Редактор

Т.А. Гладкова

Выпускающий редактор

Л.В. Петракова

Издатель

ООО «Сенсидат-Плюс»

Адрес редакции
117997, ГСП-7, Москва,
ул. Профсоюзная, д. 65, к. 410.
Тел./факс (495) 334-92-00
E-mail: pu@ipu.ru
Интернет: http://pu.mtas.ru

Оригинал-макет и электронная версия
подготовлены
ИП Прохоров О. В.

Отпечатано в ООО «Авансд солишнз»

Заказ № PB318

Подписано в печать
14.05.2018 г.

Журнал зарегистрирован
в Министерстве Российской
Федерации по делам печати,
телерадиовещания и средств
массовых коммуникаций

Свидетельство о регистрации
ПИ № ФС 77-49203 от 30 марта 2012 г.

Журнал входит в RSCI на платформе
Web of Science и Перечень
рецензируемых научных изданий ВАК

Подписные индексы:

80508 и **81708** в каталоге Роспечати;

38006 в объединенном каталоге
«Пресса России»

Цена свободная

© Федеральное государственное
бюджетное учреждение науки
Институт проблем управления
им. В.А. Трапезникова РАН

ПРОБЛЕМЫ УПРАВЛЕНИЯ

3.2018

СОДЕРЖАНИЕ

Анализ и синтез систем управления

Квинто Я.И., Хлебников М.В. Верхние оценки больших отклонений
в линейных системах при наличии неопределенности 2

Управление в социально-экономических системах

Клочков В.В., Рождественская С.М. Молодежные гранты как средство
управления развитием науки: анализ эффективности 8
Щепкин А.В. Противозатратные механизмы финансирования 17

Информационные технологии в управлении

Подлазов В.С. Бесконфликтная самомаршрутизация для трехмерного
обобщенного гиперкуба 26

Ведешенков В.А. Метод построения контролирующего теста
для цифровых систем со структурой минимального квазиполного графа
(на примере графа размера 11×11) 33

Сомов С.К. Создание информационно-технологического резерва
в распределенных системах обработки данных 40

Бабаев Д.И., Полетыкин А.Г., Промыслов В.Г., Тимофеев М.Ю.
Управление архитектурой кибербезопасности АСУТП атомных
электростанций 47

Управление техническими системами и технологическими процессами

Резчиков А.Ф., Твердохлебов В.А. Метод рекуррентного
и Z-рекуррентного управления функционированием
сложной системы 56

Управление подвижными объектами и навигация

Жилякова Л.Ю., Кузнецов Н.А., Матюхин В.Г. и др. Графовая модель
распределения локомотивов для грузовых перевозок на линейном
участке железной дороги. Задача о максимальном по включению
покрытии графика 65

Хроника

Десятая международная конференция «Управление развитием
крупномасштабных систем» MLSD'2017 76

Александр Александрович Дорофеев 81

* * *

Contents and abstracts 82

ВЕРХНИЕ ОЦЕНКИ БОЛЬШИХ ОТКЛОНЕНИЙ В ЛИНЕЙНЫХ СИСТЕМАХ ПРИ НАЛИЧИИ НЕОПРЕДЕЛЕННОСТИ¹

Я.И. Квинто, М.В. Хлебников

Рассмотрена линейная динамическая система при наличии неопределенности в ее матрице. На основе техники линейных матричных неравенств получены верхние оценки отклонений в линейных динамических системах, а также исследована задача минимизации отклонений в линейных системах управления при помощи статической линейной обратной связи по состоянию. Результаты численного моделирования продемонстрировали низкую степень консерватизма полученных оценок.

Ключевые слова: линейная динамическая система, неопределенность, линейные матричные неравенства, всплеск, функция Ляпунова.

ВВЕДЕНИЕ

Как хорошо известно, большой практический интерес представляет поведение траектории $x(t)$ устойчивой линейной системы

$$\dot{x} = Ax, \quad x(t) \in \mathbb{R}^n, \quad (1)$$

с ненулевым начальным условием. В частности, очень важна величина

$$\xi(x(0)) = \max_{t \geq 0} \frac{|x(t)|}{|x(0)|},$$

представляющая собой максимальное отклонение траектории системы от нуля в переходном режиме; $|\cdot|$ — некоторая векторная норма.

Эффект отклонений решений линейных обыкновенных дифференциальных уравнений при ненулевых начальных условиях хорошо известен и в вычислительной математике. Поскольку решение системы (1) имеет вид

$$x(t) = e^{At}x(0),$$

то

$$\xi \doteq \max_{t \geq 0} \max_{|x(0)|=1} |x(t)| = \max_{t \geq 0} \|e^{At}\|,$$

и значения отклонения непосредственно связано с оценкой матричной экспоненты; $\|\cdot\|$ — соответствующая подчиненная матричная норма.

В недавней публикации [1] были обобщены исследования больших отклонений и всплесков в линейных системах. В частности, установлены верхние оценки отклонений и предложен метод синтеза систем, гарантирующий по возможности малые отклонения.

В настоящей работе рассматриваются линейные системы при наличии неопределенности в матрице системы. На основе техники линейных матричных неравенств [2, 3] будут получены простые верхние оценки отклонений в линейных динамических системах, обладающие, как показывают результаты численного моделирования, низкой степенью консерватизма.

В § 1 путем построения общей квадратичной функции Ляпунова получены верхние оценки отклонений в линейной системе с неопределенностью (задача анализа). В § 2 будет рассмотрена задача минимизации отклонений в системах управления при помощи статической линейной обратной связи по состоянию (задача синтеза). Наконец, в § 3

¹ Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (грант № 18-08-00140).



приведен пример, иллюстрирующий предлагаемый подход.

В дальнейшем изложении все матричные неравенства понимаются в смысле знакоопределенности матриц.

1. АНАЛИЗ

Рассмотрим линейную динамическую систему

$$\dot{x} = (A + F\Delta H)x, \quad (2)$$

где $A \in \mathbb{R}^{n \times n}$, $F \in \mathbb{R}^{n \times p}$, $H \in \mathbb{R}^{q \times n}$, с фазовым состоянием $x(t) \in \mathbb{R}^n$, ненулевым начальным условием $x(0)$ и матричной неопределенностью

$$\Delta \in \mathbb{R}^{p \times q} : \|\Delta\|_2 \leq \gamma. \quad (3)$$

Матрица A предполагается гурвицевой — все ее собственные значения имеют отрицательные вещественные части.

Замечание 1. Обратим внимание на то, что матричная неопределенность Δ не предполагается фиксированной; единственное требование — ее ограниченность по норме. Таким образом, устанавливаемые далее результаты справедливы в том числе и для нестационарной неопределенности $\Delta(t)$: $\|\Delta(t)\|_2 \leq \gamma$.

В этом разделе на основе техники линейных матричных неравенств путем построения общей квадратичной функции Ляпунова будут получены верхние оценки отклонения

$$\begin{aligned} \hat{\xi} &= \max_{\|\Delta\|_2 \leq \gamma} \xi(\Delta) = \max_{\|\Delta\|_2 \leq \gamma} \max_{t \geq 0} \max_{|x(0)|_2 = 1} |x(t)|_2 = \\ &= \max_{\|\Delta\|_2 \leq \gamma} \max_{t \geq 0} \|e^{(A + F\Delta H)t}\|_2. \end{aligned}$$

Воспользуемся простым и легко проверяемым достаточным условием *робастной квадратичной устойчивости*, состоящем в наличии общей квадратичной функции Ляпунова для семейства (2). А именно, выполнение неравенства Ляпунова

$$(A + F\Delta H)P + P(A + F\Delta H)^T < 0 \quad (4)$$

с некоторой матрицей $P > 0$ при всех допустимых значениях неопределенности Δ означает, что у семейства (2) есть общая квадратичная функция Ляпунова

$$V(x) = x^T P x.$$

Далее нам понадобится вспомогательный результат, известный под названием *лемма Питерсена* [4], который эффективно применяется в разнообразных робастных постановках задач стабили-

зации и управления. Приведем его в следующей формулировке.

Лемма 1 (I.R. Petersen). Пусть $G = G^T \in \mathbb{R}^{n \times m}$, $M \in \mathbb{R}^{n \times p}$, $N \in \mathbb{R}^{q \times n}$. Матричное неравенство

$$G + M\Delta N + N^T \Delta^T M^T < 0$$

справедливо для всех $\Delta \in \mathbb{R}^{p \times q}$: $\|\Delta\|_2 \leq 1$ тогда и только тогда, когда существует число ε такое, что

$$\begin{pmatrix} G + \varepsilon M M^T & N^T \\ N & -\varepsilon I \end{pmatrix} < 0. \quad \blacklozenge$$

Таким образом, лемма Питерсена сводит проверку знакоопределенности матричного семейства

$$G + M\Delta N + N^T \Delta^T M^T$$

с матричной неопределенностью Δ к гораздо более простой задаче разрешимости линейного матричного неравенства относительно одной скалярной переменной ε .

Воспользовавшись леммой Питерсена, представим матричное неравенство (4) в виде эквивалентного ему линейного матричного неравенства относительно скалярной переменной ε и матричной переменной $P > 0$:

$$\begin{pmatrix} AP + PA^T + \varepsilon \gamma^2 FF^T & PH^T \\ HP & -\varepsilon I \end{pmatrix} < 0. \quad (5)$$

Если неравенство (5) разрешимо, то семейство (2) робастно квадратично устойчиво, и наоборот.

Заметим, что в рамках рассматриваемого подхода нетрудно вычислить радиус квадратичной устойчивости семейства (2):

$$\gamma_{\max} = \sup\{\gamma : (A + F\Delta H)P + P(A + F\Delta H)^T < 0 \text{ при некотором } P \text{ и всех } \Delta : \|\Delta\| \leq \gamma\},$$

т. е. максимальный размах $\gamma_{\max}$ неопределенности Δ такой, что при всех $\gamma < \gamma_{\max}$ у семейства (2) имеется общая квадратичная функция Ляпунова. Приведем соответствующий результат.

Лемма 2 [3]. Пусть $\hat{\rho}$ — решение задачи полуопределенного программирования

max ρ при ограничениях

$$\begin{pmatrix} AP + PA^T + \rho FF^T & PH^T \\ HP & -I \end{pmatrix} < 0, \quad P > 0,$$

относительно матричной переменной $P = P^T \in \mathbb{R}^{n \times n}$ и скалярной переменной ρ . Тогда радиус квадратичной устойчивости семейства (2)

$$\gamma_{\max} = \sqrt{\rho}. \quad \blacklozenge$$

Вернемся к задаче оценивания свеху отклонения ξ и рассмотрим эллипсоид

$$\mathcal{E} = \{x \in \mathbb{R}^n : x^T P^{-1} x \leq 1\}$$

с матрицей P , удовлетворяющей неравенству (5). Если начальное условие $x(0)$ системы (2) лежит в этом эллипсоиде, ее траектория будет оставаться в нем для всех моментов времени; это следует из того, что квадратичная форма $x^T P^{-1} x$ является функцией Ляпунова для данной системы. Следовательно, если эллипсоид $\mathcal{E}$ содержит единичный шар

$$\mathcal{B} = \{x \in \mathbb{R}^n : |x|_2 \leq 1\},$$

то для любого начального условия из шара $\mathcal{B}$ траектория системы не покинет эллипсоид $\mathcal{E}$, и в каждый момент времени для ее 2-нормы верна оценка

$$|x(t)|_2 \leq \lambda_{\max}(P) = \sqrt{\|P\|_2}.$$

Поскольку условие $\mathcal{B} \subseteq \mathcal{E}$ эквивалентно требованию $P \succcurlyeq I$, приходим к задаче

$\min \|P\|_2$ при ограничениях

$$\begin{pmatrix} AP + PA^T + \varepsilon \gamma^2 FF^T & PH^T \\ HP & -\varepsilon I \end{pmatrix} < 0, \quad P \succcurlyeq I,$$

где переменными являются матрица $P = P^T \in \mathbb{R}^{n \times n}$ и скаляр ε .

Итак, установлен следующий результат.

Теорема 1. Пусть $\hat{P}$ — решение задачи выпуклой оптимизации $\min \|P\|_2$ при ограничениях

$$\begin{pmatrix} AP + PA^T + \varepsilon \gamma^2 FF^T & PH^T \\ HP & -\varepsilon I \end{pmatrix} < 0, \quad P \succcurlyeq I,$$

относительно матричной переменной $P = P^T \in \mathbb{R}^{n \times n}$ и скалярной переменной ε .

Тогда для решений системы (2) при всех допустимых неопределенностях Δ справедлива оценка отклонения

$$\hat{\xi} \leq \sqrt{\|\hat{P}\|_2}. \quad \blacklozenge$$

Пример 1. Рассмотрим систему вида (2) с фробениусовой матрицей

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -4 & -6 & -4 \end{pmatrix},$$

обладающей собственными значениями

$$\lambda_1 = -2, \quad \lambda_{2,3} = -1 \pm j,$$

с неопределенностью

$$\Delta = (\delta_1 \delta_2 \delta_3), \quad \|\Delta\| \leq 1,$$

и «обрамляющими» матрицами

$$F = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad H = I.$$

Таким образом, возмущению подвержены элементы последней строки номинальной матрицы A .

Решение задачи из теоремы 1 дает верхнюю оценку

$$\hat{\xi} = 2,3008,$$

тогда как результат численной оптимизации

$$\max_{\|\Delta\|_2 \leq 1} \max_{t \geq 0} \|e^{(A + F\Delta H)t}\|_2$$

дает истинную величину отклонения $\xi = 1,9247$. Таким образом, полученная оценка завышена всего лишь на 16 %.

Динамика изменения величины $\max_{\|\Delta\|_2 \leq 1} \max_{|x(0)|_2 = 1} |x(t)|_2$

$= \max_{\|\Delta\|_2 \leq 1} \|e^{(A + F\Delta H)t}\|_2$ показана на рис. 1 сплошной линией.

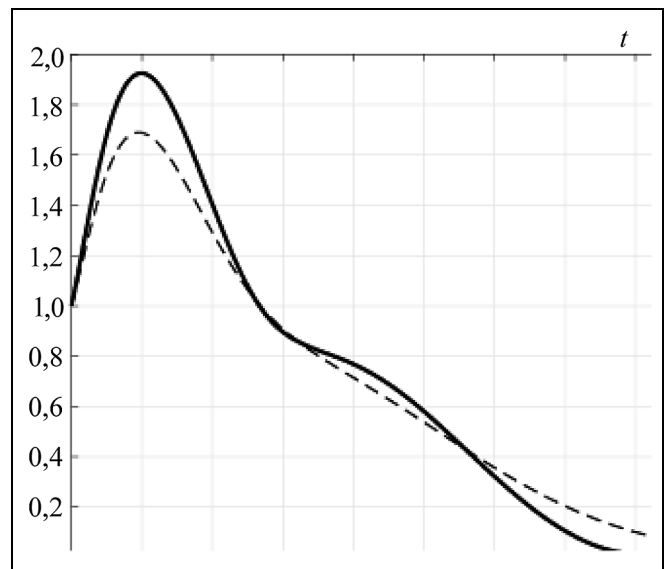


Рис. 1. График функции $\max_{\|\Delta\|_2 \leq 1} \|e^{(A + F\Delta H)t}\|_2$ для системы из примера 1



ей. Для сравнения пунктиром показан график функции $\|e^{At}\|_2$, соответствующей системе без неопределенностей.

Численное моделирование проводилось в среде Matlab с помощью программного пакета svh [5]. ♦

Оказывается, что для матриц, находящихся близко к границе устойчивости, предложенная оценка может быть сколь угодно точна.

2. СИНТЕЗ

Перейдем к задаче синтеза и рассмотрим систему управления

$$\dot{x} = (A + F\Delta H)x + Bu, \quad (6)$$

где $A \in \mathbb{R}^{n \times n}$, $F \in \mathbb{R}^{n \times p}$, $H \in \mathbb{R}^{q \times n}$, $B \in \mathbb{R}^{n \times m}$, с фазовым состоянием $x(t) \in \mathbb{R}^n$, ненулевым начальным условием $x(0)$, управлением $u(t) \in \mathbb{R}^m$ и матричной неопределенностью Δ , удовлетворяющей условию (3); пара (A, B) управляема.

Будем искать стабилизирующую линейную обратную связь по состоянию

$$u = Kx \quad (7)$$

и квадратичную функцию Ляпунова для замкнутой системы такую, что ее матрица минимальна по спектральной норме. Иными словами, в соответствии с теоремой 1 будем минимизировать оценку отклонения $\hat{\xi}$ траекторий замкнутой системы.

Замкнув систему (6) управлением (7), приходим к замкнутой системе

$$\dot{x} = (A + BK + F\Delta H)x.$$

Воспользовавшись результатами § 1, приходим к задаче $\min \|P\|_2$ при ограничениях

$$\begin{pmatrix} (A + BK)P + P(A + BK)^T + \varepsilon\gamma^2 FF^T PH^T \\ HP \quad -\varepsilon I \end{pmatrix} < 0, \\ P \succcurlyeq I. \end{pmatrix}$$

Введем вспомогательную матричную переменную $Y = KP$, тогда первое из ограничений примет линейный вид

$$\begin{pmatrix} AP + PA^T + BY + Y^T B^T + \varepsilon\gamma^2 FF^T PH^T \\ HP \quad -\varepsilon I \end{pmatrix} < 0;$$

при этом матрица регулятора K восстанавливается единственным образом: $K = YP^{-1}$.

Итак, получено следующее утверждение.

Теорема 2. Пусть $\hat{P}$, $\hat{Y}$ — решение задачи выпуклой оптимизации

$$\begin{aligned} &\min \|P\|_2 \text{ при} \\ &\begin{pmatrix} AP + PA^T + BY + Y^T B^T + \varepsilon\gamma^2 FF^T PH^T \\ HP \quad -\varepsilon I \end{pmatrix} < 0, \\ &P \succcurlyeq I, \end{aligned} \quad (8)$$

относительно матричных переменных $P = P^T \in \mathbb{R}^{n \times n}$ и $Y \in \mathbb{R}^{m \times n}$ и скалярной переменной ε .

Тогда для решений системы (6), замкнутой регулятором

$$u = \hat{K}x, \quad \hat{K} = \hat{Y}\hat{K}^{-1},$$

справедлива оценка отклонения

$$\hat{\xi} \leq \sqrt{\|\hat{P}\|_2}$$

при всех допустимых неопределенностях Δ . ♦

Важно отметить, что для системы (6) в канонической форме, т. е. при

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

с «обрамляющими» матрицами

$$F = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad H = I$$

и неопределенностью $\Delta = (\delta_0 \delta_1 \dots \delta_{n-1})$ матрица системы, замкнутой регулятором $K = (k_0 k_1 k_2 \dots k_{n-1})$, имеет вид

$$\begin{aligned} &A_c = \\ &= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ k_0 - a_0 + \delta_0 & k_1 - a_1 + \delta_1 & k_2 - a_2 + \delta_2 & \dots & k_{n-1} - a_{n-1} + \delta_{n-1} \end{pmatrix}. \end{aligned}$$

Следовательно, от коэффициентов $a_0, \dots, a_{n-1}$ и уровня γ зависит только регулятор K , тогда как матрица P функции Ляпунова и соответственно искомая оценка отклонения $\hat{\xi}$ определяются лишь

Зависимость величины $\hat{\xi}$ от n

n	2	3	4	5	6	7	8	9	10
$\hat{\xi}$	1	1,7321	2,4142	4,3357	6,7040	12,1267	19,6107	35,4989	58,8206

размерностью n системы. При этом регулятор K имеет вид

$$K = K_0 + (a_0 \ a_1 \ a_2 \ \dots \ a_{n-1}),$$

где регулятор K_0 соответствует фробениусовой матрице A с нулевой последней строкой.

Зависимость оценки отклонения $\hat{\xi}$, предоставляемой теоремой 2, от размерности n системы представлена в таблице. Отметим, что при $n = 2$ всплеск в системе отсутствует.

Замечание 2. Нетрудно видеть, что замена первого из ограничений в оптимизационной задаче (8) на условие

$$\begin{pmatrix} AP + PA^T + BY + Y^T B^T + \varepsilon \gamma^2 FF^T + 2\sigma P PH^T \\ HP \\ -\varepsilon I \end{pmatrix} < 0, \quad \sigma > 0,$$

гарантирует устойчивость матрицы $A + BK + F\Delta H + \sigma I$. Иными словами, степень устойчивости замкнутой системы (6) будет заведомо не меньше σ .

3. ПРИМЕР

Проиллюстрируем предлагаемый подход на примере так называемой *двухмассовой системы* — системы из двух твердых тел с массами m_1 и m_2 , соединенных пружиной с коэффициентом упругости k , скользящих без трения вдоль неподвижного горизонтального стержня [6]; к левому телу приложено управляющее воздействие $u \in \mathbb{R}$ (рис. 2).

Пусть x_1 и v_1 — соответственно координата и скорость левого тела, а x_2 и v_2 — правого. Поведение системы описывается вектором фазового состояния

$$x = \begin{pmatrix} x_1 \\ x_2 \\ v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^4,$$

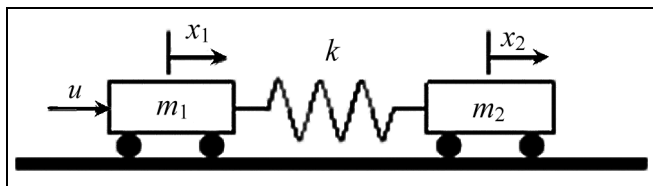


Рис. 2. Двухмассовая система

а непрерывная модель колебаний двухмассовой системы принимает вид

$$\dot{x} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k}{m_1} & \frac{k}{m_1} & 0 & 0 \\ \frac{k}{m_2} & -\frac{k}{m_1} & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{m_1} \\ 0 \end{pmatrix} u.$$

Пусть $k = 1$, $m_1 = 1$, а неопределенность системы сосредоточена в массе m_2 правого тела, на которое непосредственно не воздействует управление. При этом номинальное значение m_2 равно единице:

$$\frac{k}{m_2} = 1 + \gamma \Delta(t), \quad |\Delta| \leq 1.$$

Система, замкнутая регулятором (7), принимает вид

$$\dot{x} = (A + BK + F\Delta H)x,$$

где

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \gamma \end{pmatrix}, \quad H = (1 \ -1 \ 0 \ 0).$$

Воспользовавшись теоремой 2, для допустимого уровня возмущения $\gamma = 0,85$ находим матрицу

$$\hat{P} = \begin{pmatrix} 32,9119 & 29,7318 & -3,1863 & -5,6495 \\ 29,7318 & 34,4152 & 7,7303 & -3,9182 \\ -3,1863 & 7,7303 & 62,3404 & -0,9822 \\ -5,6495 & -3,9182 & -0,9822 & 2,7200 \end{pmatrix}$$

квадратичной функции Ляпунова и соответствующий регулятор

$$\hat{K} = (-11,0638 \ 8,3672 \ -2,3681 \ -10,4377);$$

при этом получаем верхнюю оценку отклонения $\hat{\xi} = 8,1927$.

На рис. 3 показана динамика величины $\|x(t)\|$ (жирная линия) и траектории фазовых координат (тонкие линии) замкнутой системы при некотором начальном условии $x(0)$ из единичного шара. В качестве матричной неопределенности примем так называемую «наихудшую неопределенность»:

$$\tilde{\Delta} = \text{sign}[F^T \hat{P}^{-1} x x^T H^T].$$

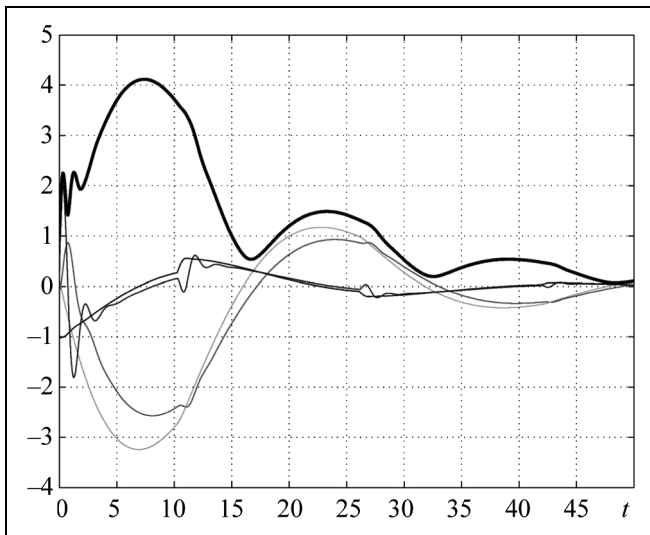


Рис. 3. Всплеск в двухмассовой системе

Как показано в работе [7], матричная неопределенность $\tilde{\Delta}$ (рассматриваемая как функция от времени) максимизирует производную функции Ляпунова в силу системы и в этом смысле является «наихудшей» из всех допустимых неопределенностей.

ЗАКЛЮЧЕНИЕ

Исследования переходных режимов в линейных динамических системах при ненулевых начальных условиях были начаты еще в 1948 г. в работе А.А. Фельдбаума [8]. В настоящей статье продолжено изучение этого явления для линейных систем при наличии структурированной матричной неопределенности в матрице системы. На основе техники линейных матричных неравенств получены верхние оценки отклонений в линейных динамических системах, а также исследована задача минимизации отклонений в линейных системах управления при помощи статической линейной обратной связи по состоянию. Результаты численного моделирования демонстрируют низкую степень консерватизма полученных оценок.

Полученные результаты авторы в дальнейшем предполагают распространить на иные робастные постановки задачи, на системы в дискретном времени, а также на динамические системы, подверженные воздействию внешних возмущений.

ЛИТЕРАТУРА

1. Поляк Б.Т., Тремба А.А., Хлебников М.В. и др. Большие отклонения в линейных системах при ненулевых начальных условиях // Автоматика и телемеханика. — 2015. — № 6. — С. 18—41 [Polyak B.T., Tremba A.A., Khlebnikov M.V., et al. Large Deviations in Linear Control Systems with Nonzero Initial Conditions // Automation and Remote Control. — 2015. — Vol. 76, N 6. — P. 957—976].
2. Boyd S., El Ghaoui L., Feron E., Balakrishnan V. Linear Matrix Inequalities in Systems and Control Theory. — Philadelphia: SIAM, 1994.
3. Поляк Б.Т., Хлебников М.В., Шербаков П.С. Управление линейными системами при внешних возмущениях: Техника линейных матричных неравенств. — М.: ЛЕНАНД, 2014. — 560 с.
4. Petersen I.R. A Stabilization Algorithm for a Class of Uncertain Linear Systems // Systems and Control Letters. — 1987. — Vol. 8. — P. 351—357.
5. Grant M., Boyd S. CVX: Matlab Software for Disciplined Convex Programming, Version 1.21. — URL: <http://cvxr.com/cvx/>.
6. Reinelt W. Robust Control of a Two-Mass-Spring System Subject to Its Input Constraints // Proc. American Control Conference — 2000 (ACC-2000), Chicago, USA, June 28—30, 2000. — P. 1817—1821.
7. Поляк Б.Т., Топунов М.В., Шербаков П.С. Идеология инвариантных эллипсоидов в задаче о робастном подавлении ограниченных внешних возмущений // Стохастическая оптимизация в информатике. — 2007. — Т. 3, № 1-1. — С. 51—84.
8. Фельдбаум А.А. О распределении корней характеристического уравнения систем регулирования // Автоматика и телемеханика. — 1948. — № 4. — С. 253—279.

Статья представлена к публикации членом редколлегии С.А. Красновой.

Квинто Яна Игоревна — канд. техн. наук, вед. инженер,
✉ yanakvinto@mail.ru,

Хлебников Михаил Владимирович — д-р. физ.-мат. наук,
гл. науч. сотрудник, ✉ khlebnik@ipu.ru,

Институт проблем управления им. В.А. Трапезникова РАН,
г. Москва.

Не забудьте подписаться!

Подписку на журнал «Проблемы управления» можно оформить в любом почтовом отделении (подписной индекс 81708 в каталоге Роспечати или 38006 в объединенном каталоге «Пресса России»), а также через редакцию с любого месяца, при этом почтовые расходы редакция берет на себя. Отдельные номера редакция высылает по первому требованию.

МОЛОДЕЖНЫЕ ГРАНТЫ КАК СРЕДСТВО УПРАВЛЕНИЯ РАЗВИТИЕМ НАУКИ: АНАЛИЗ ЭФФЕКТИВНОСТИ¹

В.В. Клочков, С.М. Рождественская

Проведен анализ влияния финансовой поддержки в виде молодежных грантов на эффективность отбора более способных ученых. Оценено влияние финансовой обеспеченности ученого на время, которое он может посвящать науке, его результативность в период грантовой поддержки и, в конечном итоге, на качество отбора ученых по итогам их ранжирования. Определены требования к уровню грантовой поддержки. Оценено влияние на качество отбора длительности периода поддержки и наблюдения за результативностью ученых. Определены желательные временные и стоимостные параметры грантовой поддержки молодых ученых.

Ключевые слова: управление развитием, гранты, молодые ученые, жизненный цикл, продуктивность, ранжирование, эффективность, ошибки первого и второго рода.

ВВЕДЕНИЕ

Традиционно принято оказывать особую поддержку молодым научным работникам. Считается крайне важным привлечь молодых специалистов в науку, чтобы обеспечить воспроизводство кадров и преемственность поколений в этой сфере, что особенно актуально ввиду тяжелых кадровых потерь отечественной науки с начала постсоветского периода. Однако такой мотив «поддержки молодежи» не уникален для науки и актуален для многих видов деятельности, особенно для высокотехнологичной промышленности. Предположим, что нынешние кадровые проблемы «переходного периода» будут преодолены. Нужно ли будет и далее «поддерживать молодых ученых»?

Нередко, особенно при сокращении ресурсных возможностей для финансирования научной сферы, возникает вопрос: нужны ли вообще механизмы финансовой поддержки научной молодежи и, в частности, особые гранты для молодых ученых, какова реальная польза таких механизмов финансирования для развития российской науки в целом (а не только для их получателей)? В ряде работ, например, [1–3], проводится, скорее, качественный анализ достоинств и недостатков грантовой системы финансирования научных исследований, иногда — статистический анализ апостериорной

результативности работы грантовых фондов. В настоящей работе, в отличие от предыдущих, проводится экономико-математическое моделирование возможного механизма влияния молодежных грантов на развитие науки, а также анализ возможностей системы молодежных грантов как инструмента управления развитием науки.

1. ОСНОВНАЯ ГИПОТЕЗА ИССЛЕДОВАНИЯ: ФИНАНСОВАЯ ПОДДЕРЖКА УЧЕНЫХ КАК СРЕДСТВО ПОВЫШЕНИЯ КАЧЕСТВА ИХ ОТБОРА

Молодые исследователи способны вносить ценный вклад в науку, в том числе выдвигая смелые и нестандартные идеи, и необходимо обеспечить им возможности для этого. Но в данной работе в центре внимания — иные аспекты материальной поддержки молодых ученых. Если различные формы базовой поддержки молодых исследователей призваны обеспечить устойчивое воспроизводство кадров, то гранты служат формой конкурсного финансирования. И хотя отдельный конкурс для молодежи призван создать ей более благоприятные условия, однако и в таком конкурсе будут победители и проигравшие. Поэтому молодежные гранты — это уже попытка отобрать в общем потоке молодых специалистов, приходящих в науку, более перспективных, по тем или иным критериям.

Как правило, конкуренция среди ученых весьма жесткая. И на протяжении жизненного цикла претенденты на право заниматься наукой постепенно отсеиваются, уходя в другие виды деятельности.

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 17-32-00025).

Механизмы финансирования научных исследований в значительной степени также конкурентные (это касается не только грантового финансирования — ведь и базовое финансирование выделяется на те научные позиции, которые сами по себе являются предметом конкуренции). Современная организация профессиональной науки предполагает ранжирование ученых, выделение более квалифицированных или заслуженных и отсеивание проигравших в этой конкуренции. В работах авторов и других исследователей [4–6] ставится под сомнение целесообразность ужесточения конкуренции в научной сфере. Так или иначе, внешняя оценка квалификации ученого и его заслуг неизбежна, но нужно учитывать возможность ошибок при таком оценивании и определять их влияние на эффективность науки в целом, на перспективы ее развития. Авторы работы [5] предложили простую модель ранжирования ученых по некоторым формальным признакам. Ранжирование проводится с целью отделения «лидеров», «активно работающих и талантливых ученых», от «научного балласта» для того, чтобы в дальнейшем именно первой категории ученых оказывать поддержку, а второй предложить уйти из профессии. По мнению апологетов конкурентной организации науки, это позволит повысить ее эффективность, измеряемую как соотношение затрат на ученых и их продуктивности. В свою очередь, результативность научной деятельности может измеряться какими-либо формальными показателями — например, числом публикаций (в том числе в определенных изданиях), их цитируемостью и т. п. В предложенной простейшей модели ученые делятся на два класса — высокопродуктивные в описанном смысле и низкопродуктивные (лидеры и аутсайдеры). Учитываются возможности ошибок двух видов при ранжировании (используется терминология, традиционная для статистических методов проверки гипотез):

- ошибки первого рода (отсеять истинно высокопродуктивного ученого), вероятность которой обозначим $(1 - p)$;
- ошибки второго рода (признать высокопродуктивным посредственного исследователя), вероятность которой обозначим q .

Предполагается, что средняя производительность лидера в x раз выше по сравнению с аутсайдерами. Тогда, если истинная доля лидеров в научном сообществе составляет α , средняя произво-

Характеристики формальных критериев ранжирования ученых (пример)

№ варианта	1	2	3
p	0,8	0,6	0,9
q	0,4	0,2	0,1

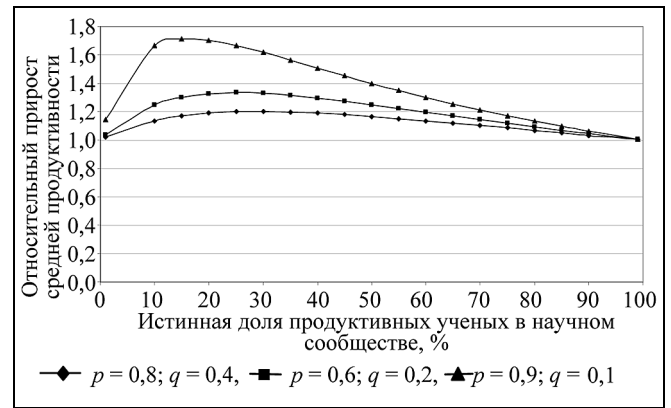


Рис. 1. Зависимость относительного изменения средней производительности ученых вследствие их отбора по формальным критериям ($x = 3$)

дительность ученых, отобранных по критериям с указанными выше характеристиками (вероятностями ошибок первого и второго рода) возрастет в

$$\frac{\bar{u}'}{\bar{u}} = \frac{p_{\text{апост}}^1 x + (1 - p_{\text{апост}}^1)}{\alpha x + (1 - \alpha)} \quad (1)$$

раз, где $p_{\text{апост}}^1$ — апостериорная вероятность того, что отобранный ученый в самом деле оказался лидером. Эта вероятность определяется с помощью формулы Байеса:

$$p_{\text{апост}}^1 = \frac{p\alpha}{p\alpha + q(1 - \alpha)}. \quad (2)$$

В таблице рассмотрено три варианта формальных критериев классификации ученых. Первый — «мягкий» критерий ($p = 0,8$; $q = 0,4$) — достаточно хорошо выделяет сильных ученых, но не очень хорошо отсеивает слабых. Второй — «жесткий» критерий ($p = 0,6$; $q = 0,2$) — наоборот, достаточно часто ошибается при выделении сильных ученых, но лучше отсеивает слабых. Третий критерий ($p = 0,9$; $q = 0,1$) гораздо более точен и надежен, чем первые два, в обоих отношениях — вероятности ошибок первого и второго рода одинаковы и составляют лишь 10 %.

На рис. 1 и 2 изображены полученные по вышеприведенным аналитическим формулам (1) и (2) графики зависимости относительного изменения средней продуктивности ученых (измеряемой в условных «результатах» за период — публикациях, ссылках на них и т. п.) в результате отбора по формальным критериям № 1–3 (см. таблицу), от фактической доли высокопродуктивных ученых в научном сообществе α . При этом считается, что средняя производительность лидера, измеренная в описанных условных результатах в год, в $x = 3$ и в $x = 10$ раз выше, чем производительность «посредственного» ученого. Графики имеют максимум

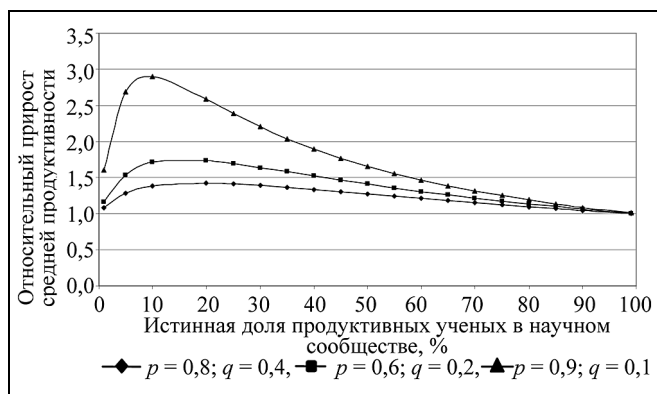


Рис. 2. Зависимость относительного изменения средней производительности ученых вследствие их отбора по формальным критериям ($x = 10$)

при некоторых значениях фактической доли высокопродуктивных ученых в научном сообществе α . Таким образом, вне зависимости от этой доли — в реальности, ненаблюдаемой — можно оценить сверху максимально достижимую эффективность ранжирования ученых. При трехкратном превосходстве лидеров над «посредственными» учеными (см. рис. 1) максимальный прирост средней эффективности научной работы после отбора ученых по формальным показателям составит лишь 20—37 % соответственно. И даже относительно точный критерий № 3 при трехкратном превосходстве лидеров обеспечит прирост средней производительности труда, самое большее, лишь на 73 %. Несколько выше эффективность ранжирования ученых, если преимущество лидеров над «посредственными» коллегами будет десятикратным (рис. 2). Однако даже тогда прирост средней производительности ученых составит для грубых критериев, самое большее, 40—80 %, а для наиболее точного критерия будет менее чем трехкратным (тогда как, напомним, выделяемые по этому критерию лидеры превосходят своих коллег в истинной производительности в 10 раз).

Таким образом, как показано в работе [5], при достижимых в реальности показателях достоверности библиометрических критериев ранжирования (при вероятностях ошибок первого и второго рода порядка 0,1 и более) ожидаемый прирост средней эффективности ученых после такого ранжирования будет невелик. Подчеркнем, что сама по себе средняя продуктивность ученых — чрезвычайно примитивный критерий эффективности управления наукой. Но в статье [5] авторы приняли именно его, чтобы проверить — достигают ли апологеты формальных библиометрических критериев хотя бы заявляемых ими целей «повышения средней отдачи от науки на вложенный рубль», даже если отдачу понимать именно как поток пуб-

ликаций (безотносительно к их истинной значимости) или их цитируемости. При этом игнорируется более сложный, в реальности, коллективный характер научной работы и самого научного знания, как подчеркивалось в заключение статьи [5].

В настоящей работе используется тот же критерий эффективности науки, поскольку рассматривается механизм влияния на него молодежных грантов, применяемых в сложившейся «конкурентной» системе управления и финансирования науки. Что влияет на достоверность измерения результативности ученых, на вероятности ошибок ранжирования? Молодые исследователи, только начинающие свой путь в науке, как правило, не могут полностью обеспечивать себя за счет научной работы (особенно в странах, испытывающих дефицит средств на финансирование науки) и вынуждены подрабатывать более или менее далекими от области своих научных интересов способами. Причем это касается даже ведущих экономически развитых стран мира, где молодые исследователи вынуждены преподавать в университетах значительную часть своего времени, проводя семинарские занятия и ассистируя своему научному руководителю. Вопреки стереотипу, далеко не всегда такая преподавательская деятельность напрямую связана с тематикой научной работы аспиранта или постдока. В работе авторов [7] предложена простая экономико-математическая модель результативности ученого в зависимости, прежде всего, от его индивидуальной производительности (продуктивности), а также от уровня финансирования научной работы (она, в свою очередь, опирается на общие идеи в области моделирования индивидуального предложения труда, изложенные в работах [8, 9]). Учитываются такие факторы, как:

- индивидуальная потребность в средствах (которая как раз у молодых работников любой профессии выше, поскольку они должны создавать семью, растить детей и в то же время обзаводиться жильем, тогда как у более зрелых работников основная доля «инвестиционных расходов», как правило, уже позади);

- ставка оплаты «подработки», которая у более зрелых ученых, как правило, выше, чем у молодых, поскольку их статус и репутация позволяют получать и более выгодные предложения дополнительного заработка.

Качественные выводы из анализа построенной модели таковы. Молодые ученые, особенно в условиях, когда базовый уровень финансирования науки низок, вынуждены посвящать собственно научным исследованиям лишь очень малую часть своего рабочего времени. Потому их результативность лишь в малой степени отражает их реальные способности, их истинную научную продуктивность и, тем более, потенциал роста.

Таким образом, измеримая результативность работы ученых может сильно отличаться от их реальной продуктивности, особенно для молодых ученых, вынужденных тратить значительное время на подработку. Потому необходимо исследовать влияние этих факторов на точность и достоверность «оценки со стороны» продуктивности молодых исследователей. Описанные особенности ранних стадий жизненного цикла ученых таковы, что истинная продуктивность молодого ученого, его потенциал, выявляются лишь в малой степени. Гораздо большее влияние на измеримые показатели результативности ученого оказывают иные факторы — доступные возможности подработки, личные обстоятельства, повышающие индивидуальные потребности в средствах, и др. Причем они индивидуальны не в меньшей степени, чем собственно способности к науке (в том числе и по причине разницы в происхождении, финансовых возможностях семьи и др.). В итоге весьма вероятно, что истинно высокопродуктивный, способный ученый на ранних стадиях покажет относительно низкую результативность, тогда как менее способный — более высокую; т. е. «молодежные», точнее — «стартовые» эффекты резко повышают вероятности ошибок первого и второго рода при ранжировании ученых, что в дальнейшем снижает эффективность науки в целом. Если истинные способности молодых ученых выявляются недостоверно, с одной стороны, отсеиваются и уходят из науки более способные к ней, а с другой стороны — получают поддержку и далее остаются в профессии менее способные. Поэтому молодежные гранты (а также прочие формы материальной поддержки), будучи некоторым авансом молодым исследователям, избавляют их — в большей или меньшей степени — от необходимости подработки, позволяют сосредоточиться на научной работе, продемонстрировать свою истинную продуктивность. Далее возрастают точность оценивания продуктивности ученых, достоверность их ранжирования и отсева менее способных и, в конечном итоге — средняя продуктивность тех исследователей, которые смогут остаться в профессии и продолжить научную карьеру. В данной работе эти качественные предположения проверяются путем математического моделирования. Затем на основе анализа построенных моделей будут выработаны некоторые рекомендации по совершенствованию механизмов материальной поддержки молодых ученых.

2. МОДЕЛЬ ВЛИЯНИЯ ФИНАНСОВОЙ ПОДДЕРЖКИ НА ДОСТОВЕРНОСТЬ ИЗМЕРЕНИЯ РЕЗУЛЬТАТИВНОСТИ РАБОТЫ УЧЕНЫХ

Количественную оценку эффективности описанных выше механизмов финансовой поддержки

ученых можно получить следующим образом. Рассмотрим период наблюдения над результативностью ученого длительностью t лет (по итогам этого периода и будет приниматься решение о его дальнейшей профессиональной судьбе). За этот период ученый может получить некоторое случайное количество обобщенных «результатов» (публикаций в определенных изданиях, ссылок на эти публикации и т. п.). Они и служат мерилем продуктивности ученого. И хотя их количество случайно, но интенсивность их появления прямо пропорциональна, прежде всего, уровню индивидуальной продуктивности ученого, отражающей его способности, а также времени, которое ученый может уделять собственно науке, а не подработке. За t лет это время составит $t\delta_{\text{НИР}}$, где $\delta_{\text{НИР}}$ — доля календарного фонда рабочего времени, которую ученый может посвятить именно научной работе. Согласно модели, предложенной авторами в работе [7], эта доля выражается формулой:

$$\delta_{\text{НИР}}(k_0) = 1 - k_{\text{пут}}(1 - k_0), \quad (3)$$

где $k_0 = z_0/z_{\min}$ — коэффициент достаточности базового финансирования (в свою очередь, z_0 , ден. ед./год — уровень базового финансирования научной работы, при полной занятости; $z_{\min}$ — минимально необходимый данному ученому уровень

годового дохода), $k_{\text{пут}} = \frac{z_{\min}}{w \cdot 2912}$ — отношение минимально необходимого дохода к оплате подработки (здесь ставка оплаты подработки обозначена w , ден. ед./ч, а 2912 ч — календарный годовой фонд рабочего времени при полной занятости).

Разумеется, должно выполняться неравенство $z_0 \leq z_{\min} \leq z_0 + 2912w$, гарантирующее, что доля времени на НИР лежит в пределах от 0 до 1.

Именно описанную долю рабочего времени $\delta_{\text{НИР}}$ и помогает повысить финансовая поддержка ученого, приближая z_0 к $z_{\min}$.

Пусть продуктивность «посредственного» ученого, т. е. интенсивность появления результатов в течение года при 100%-й занятости научной работой, составляет y результатов, а для лидеров — в x раз больше, т. е. xy . Пусть закон распределения потока «результатов» является простейшим, или пуассоновским. Тогда вероятности того, что за период t ученые — лидеры и аутсайдеры — получат ровно y результатов, определяются формулами вида:

$$p_{\text{аутсайдер}}(y, t) = \frac{e^{-\bar{y}(t)} [\bar{y}(t)]^y}{y!}; \quad (4)$$

$$p_{\text{лидер}}(y, t) = \frac{e^{-x\bar{y}(t)} [x\bar{y}(t)]^y}{y!}; \quad (5)$$

где

$$\bar{y}(t) = ut\delta_{\text{НИР}} \quad (6)$$

— среднее число результатов аутсайдера за период наблюдения t .

Далее на основании измеренного случайного числа результатов ученых по итогам ранжирования будет отнесен к высокопродуктивным (и, соответственно, получит постоянную позицию, сможет остаться в науке и далее повышать свой статус и квалификацию) или к низкопродуктивным (и будет вынужден уйти из профессии). Проводящие ранжирование (органы государственного управления, фонды, финансирующие научные исследования и др.) сами устанавливают порог $y_{\text{порог}}$, отделяющий избранных ученых от прочих. При этом они не могут оптимизировать его по интегральным критериям повышения средней производительности выделенных ученых, поскольку:

— не знают истинной доли высокопродуктивных ученых в общей массе¹ α ;

— тем более, не знают индивидуальных факторов, определяющих возможности конкретного ученого заниматься собственно наукой (т. е., в терминах предлагаемых моделей, не знают доступной данному ученому ставки оплаты подработки w , а также минимально необходимого ему дохода $z_{\text{мин}}$).

Рассмотрим последствия установления порога на некотором произвольном уровне $y_{\text{порог}}$. Вероятности того, что такой порог преодолит высокопродуктивный, т. е. способный ученый, а также ученый с низкой продуктивностью, можно выразить через интегральные функции распределения Пуассона (они равны вероятностям того, что результативность обеих категорий ученых за период наблюдения не превысит установленного порога):

$$P_{\text{аутсайдер}}(y_{\text{порог}}, t) = \sum_{i=0}^{i=y_{\text{порог}}} p_{\text{аутсайдер}}(i, t); \quad (7)$$

$$P_{\text{лидер}}(y_{\text{порог}}, t) = \sum_{i=0}^{i=y_{\text{порог}}} p_{\text{лидер}}(i, t). \quad (8)$$

Фактически, эти функции и определяют вероятности ошибок первого и второго рода в модели ранжирования ученых:

$(1 - p) = P_{\text{лидер}}(y_{\text{порог}}, t)$ — вероятность ошибки первого рода, т. е. отсева высокопродуктивного ученого;

¹ Согласно выводам, полученным в работе [5], от этой доли зависит относительный прирост средней продуктивности ученых по итогам их ранжирования и отбора, а также оптимальная политика проводящих ранжирование и отбор ученых. Это наглядно показывают и графики на рис. 1 и 2.

$q = 1 - P_{\text{аутсайдер}}(y_{\text{порог}}, t)$ — вероятность ошибки второго рода, т. е. отбора низкопродуктивного ученого.

Итак, пользуясь формулами (4)–(8), можно оценить в рамках принятых здесь предпосылок вероятности преодоления порогов отбора учеными-лидерами и учеными-аутсайдерами в зависимости от управляющих переменных — от доли времени, посвящаемой науке (она растет благодаря грантовой поддержке), от жесткости самого порога и от длительности периода наблюдения (и, опять же, грантовой поддержки молодых исследователей):

$$p = 1 - P_{\text{лидер}} = p(y_{\text{порог}}, t, \delta_{\text{НИР}});$$

$$q = 1 - P_{\text{аутсайдер}} = q(y_{\text{порог}}, t, \delta_{\text{НИР}}).$$

Далее, в принципе, получившиеся выражения для вероятностей ошибок первого и второго рода можно подставить в формулы (1), (2) и оценить относительный прирост средней производительности ученых после их отбора и ранжирования: $\bar{u}'/\bar{u} = f(y_{\text{порог}}, t, \delta_{\text{НИР}})$. В свою очередь, уровень грантовой поддержки Δz можно подставить в формулу (3), строго выразив обеспечиваемую этим уровнем долю времени, посвящаемую науке:

$$\delta_{\text{НИР}}(\Delta z) = 1 - \frac{z_{\text{мин}}}{w \cdot 2912} \left(1 - \frac{z_0 + \Delta z}{z_{\text{мин}}} \right).$$

Далее остается учесть, что ограниченный бюджет M , выделяемый на грантовую поддержку совокупности из N молодых ученых, может расходоваться по-разному: можно повышать размер грантов Δz или длительность периода грантовой поддержки t :

$$\bar{u}'/\bar{u} \rightarrow \max_{y_{\text{порог}}, t, \Delta z} |\Delta z t N = M,$$

т. е. можно полностью формализовать задачу управления развитием науки в рамках описанного здесь механизма. Также можно разнообразить критерии оптимальности — прежде всего, пользуясь моделями [5], оценить сокращение численности ученых при таком отборе и снижения общей продуктивности научной сферы, неизбежного при отсеивании части ученых. Но в этой работе такой формальный анализ до конца проводить нецелесообразно. Достаточно качественно оценить вероятности преодоления установленных порогов учеными-лидерами и учеными-аутсайдерами в зависимости от доли времени, посвящаемой науке, при грантовой поддержке, от ее длительности и жесткости порогов отбора. Полученные при этом количественные оценки уже позволят обосновать качественные выводы и рекомендации.

3. ПАРАМЕТРИЧЕСКИЙ АНАЛИЗ ВЛИЯНИЯ ДОЛИ ВРЕМЕНИ, ПОСВЯЩАЕМОЙ НАУКЕ, И ЖЕСТКОСТИ ПОРОГА РАНЖИРОВАНИЯ НА КАЧЕСТВО ОТБОРА УЧЕНЫХ

При низком уровне ресурсной обеспеченности научной деятельности (и, соответственно, при малой доле времени, которую ученые смогут тратить на занятие наукой), интенсивности обоих потоков результатов будут относительно низкими, поскольку ученые вынуждены тратить большую часть времени на подработку. Как это отразится на вероятностях ошибок при ранжировании и отборе ученых?

Предположим, что при ранжировании в качестве порога выбрано:

а) ожидаемое число результатов, которое должен получить высокопродуктивный ученый за период наблюдения, т. е. $y_{\text{порог}} = x u t \delta_{\text{НИР}}$;

б) 70 % ожидаемой результативности лидера, т. е. $y_{\text{порог}} = 0,7 x u t \delta_{\text{НИР}}$;

в) 30 % ожидаемой результативности лидера, т. е. $y_{\text{порог}} = 0,3 x u t \delta_{\text{НИР}}$

(результат округляется до ближайшего целого числа, но не менее 1). Особо подчеркнем, что с ростом доли времени, уделяемой науке, $\delta_{\text{НИР}}$, и порог ранжирования возрастает, поскольку ожидается, что ученые могут больше времени посвящать науке, а не подработке. И наоборот, понимая, что ученые вынуждены подрабатывать, контролирующие органы или научные фонды снижают требования к результативности за период наблюдения. Такое допущение весьма оптимистично, поскольку для проведения описанной политики орган управления наукой должен знать реальные условия работы и жизни ученых — пусть даже не индивидуальные, а среднестатистические. На практике это не выполняется, в том числе и по причине преднамеренного искажения администрацией научных и образовательных учреждений сведений о доходах ученых¹.

Найдем вероятности того, что установленный порог преодолеют высоко- и низкопродуктивный ученые, которые в реальности посвящают научной работе произвольную долю времени $\delta_{\text{НИР}}$ в диапазоне от 20 до 100 %. На рис. 3 изображены графики

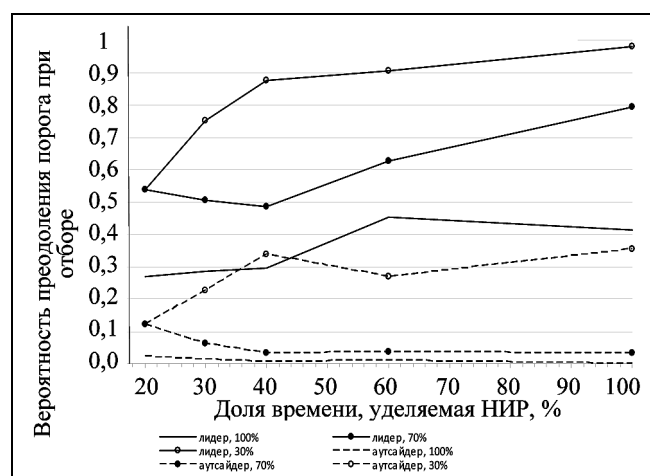


Рис. 3. Зависимость вероятности прохождения отбора учеными различной продуктивности от доли времени, посвящаемой научной работе ($t = 3$ года)

зависимости вероятностей преодоления описанных порогов ранжирования (варианты А — В) для высоко- и низкопроизводительного ученых от доли времени, посвящаемой науке, $\delta_{\text{НИР}}$. Эти графики получены аналитическим расчетом по вышеприведенным формулам (7) и (8) вероятностей непреодоления установленных порогов, в которые, в свою очередь, подставлены вероятности получения за период наблюдения ровно у результатов из формул (4) и (5). Эти вероятности зависят от доли времени, посвящаемой научной работе, согласно формуле (6). Здесь приняты такие исходные данные: $u = 1$ результат в год; $x = 3$ (т. е. лидеры превосходят аутсайдеров в продуктивности втрое); $t = 3$ года. Из этих графиков видно², что при относительно низких долях времени, уделяемого науке, $\delta_{\text{НИР}} = 20\text{—}30\%$, вероятность преодолеть любые пороги ранжирования (даже относительно мягкий, 30%-й, соответствующий варианту В) для аутсайдеров низка, т. е. отсев низкопродуктивных ученых будет весьма достоверным. В то же время, и высокопродуктивные ученые преодолеют 50%-й порог (вариант Б) с вероятностью около 0,5, а 30%-й (вариант В) — с вероятностью лишь око-

¹ Стремясь формально выполнить требования так называемых «майских» (2012 г.) указов Президента РФ, в частности, требование обеспечения средней зарплаты научных работников и преподавателей вузов на уровне не ниже 1,5—2 средних зарплат по региону, администрации переводят работников на все меньшие дробные части ставок (при том, что реальный объем работы не сокращается), обеспечивая рост зарплаты в расчете на ставку. Такая политика неизбежна, поскольку выполнение «майских» указов не было обеспечено ростом финансирования и фонда зарплаты в вузах и научных организациях.

² Следует пояснить немонокотонность графиков и наличие на них участков убывания — хотя, на первый взгляд, чем выше доля времени, посвящаемая НИР, тем выше вероятность преодоления порога отсева. Но с ростом доли времени, посвящаемой НИР, в данной модели возрастают и нормативы продуктивности. Т. е. финансирующие науку и контролирующие органы ожидают, что, посвящая науке больше времени, ученые должны получать больше результатов — и соответственно повышают планку отбора (что весьма близко к реальности). Кроме того, эта планка определяется в данной модели дискретным образом, и если ожидаемое количество результатов лидера за период будет дробным, норматив получается округлением этой дроби до ближайшего целого числа.

ло 0,7. Что касается варианта А, т. е. попытки потребовать от высокопродуктивных ученых показать результативность на уровне ее математического ожидания за период наблюдения, такой порог ученые-лидеры преодолеют с вероятностью лишь около 0,3. Отсев будет настолько жестким, что потери кадрового потенциала науки окажутся существенно более весомыми, чем повышение ее средней эффективности (согласно модели [5], оно составит, в лучшем случае, лишь 20—30 %). Таким образом, в этом случае целесообразно принять относительно мягкий, 30%-й порог ранжирования ученых, что обеспечит вероятности ошибок первого рода не более 0,25—0,45, и второго рода — не более 0,12—0,22. Однако при этом возможный прирост средней продуктивности ученых по итогам их отбора не превысит 20 %.

Если же ученые смогут уделять науке большую часть своего рабочего времени, свыше 60—70 %, тогда, как видно из графиков, вероятности ошибок первого и второго рода несколько сократятся. В данном примере высокопродуктивные ученые смогут преодолевать 30%-й порог почти гарантированно, а 70%-й — с вероятностью 0,6 и выше, т. е. вероятности ошибок первого и второго рода составят:

— для 30%-го порога (вариант В) — соответственно, не более 0,1 и 0,28 (что лучше «мягкого» критерия в таблице);

— для 70%-го порога (вариант Б) — соответственно, не более 0,37 и около 0,05 (что лучше «жесткого» критерия в таблице).

В таких условиях наибольший прирост эффективности науки (и относительно приемлемые потери количественного состава ученых) обеспечит 70%-й порог (вариант Б). Наиболее жесткий, 100%-й порог результативности, с одной стороны, обеспечивает почти гарантированный отсев аутсайдеров. С другой стороны, и отсев лидеров составит около 50 %, даже если ученые все рабочее время будут посвящать науке.

Особо подчеркнем: здесь считается, что высоко- и низкопродуктивные ученые могут уделять науке одинаковую долю времени, т. е. интенсивности потоков получаемых ими результатов отличаются во столько же раз (в пользу «лидеров»), во сколько раз «лидеры» продуктивнее «аутсайдеров». И даже при таком допущении случайные колебания числа результатов за период наблюдения приводят к тому, что достоверно отделить «лидеров» от остальных исследователей не удастся. В реальности условия, в которых проводится ранжирование ученых, будут еще более пессимистичными. Более чем возможно, что низкопродуктивный «аутсайдер», тем не менее, обладая лучшими стартовыми финансовыми возможностями, сможет уделять науке больше времени, чем «лидер», и даже по ожидаемой результативности будет превосходить

потенциального «лидера». Такие ситуации нередко встречаются в реальности. Положение усугубляется усилением имущественного неравенства, платности образования и др., ослаблением «социальных лифтов» в рыночной экономике. С учетом этих аспектов финансовая поддержка молодых ученых еще более актуальна и эффективна. Выделение молодым ученым суммы, достаточной для полноценных занятий наукой, в идеале полностью нивелирует различие в индивидуальных финансовых возможностях и потребностях, позволяя оценивать истинную продуктивность ученых.

4. ПАРАМЕТРИЧЕСКИЙ АНАЛИЗ ВЛИЯНИЯ ДЛИТЕЛЬНОСТИ ПЕРИОДА НАБЛЮДЕНИЯ НА КАЧЕСТВО РАНЖИРОВАНИЯ И ОТБОРА УЧЕНЫХ

Можно заметить, что в рассмотренном выше примере общая продолжительность периода наблюдения относительно невелика, с учетом принятых значений продуктивности ученых. Наиболее вероятные численности результатов за период наблюдения будут малыми — порядка нескольких единиц, что и порождает высокие вероятности ошибок ранжирования. На рис. 4 приведены графики, аналогичные таковым на рис. 3, но полученные для еще более короткого периода наблюдения — 1 год. В этом случае, как можно видеть на графиках, ранжирование и отбор ученых вообще теряют смысл, поскольку их достоверность будет неприемлемо низкой даже при 100%-й занятости ученых исследованиями. При высокой доле времени, посвящаемой НИР (близкой к 100 %), наиболее целесообразным будет установить мягкий порог ранжирования, соответствующий 30 % ожидаемой результативности ученых-лидеров.

Однако более рациональным было бы увеличение длительности периода наблюдения. Возмож-

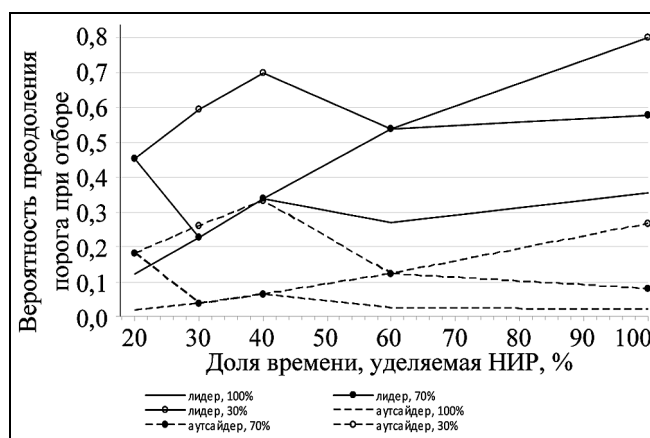


Рис. 4. Зависимость вероятности прохождения отбора учеными различной продуктивности от доли времени, посвящаемой научной работе ($t = 1$ год)

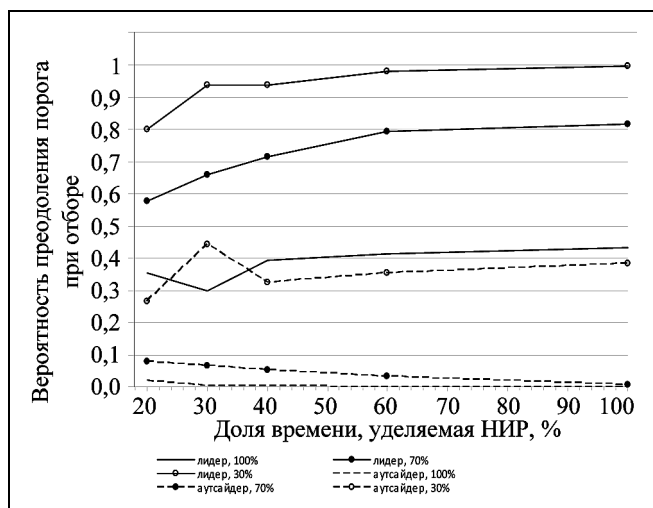


Рис. 5. Зависимость вероятности прохождения отбора учеными различной продуктивности от доли времени, посвящаемой научной работе ($t = 5$ лет)

ный прирост затрат на финансовую поддержку ученых в течение этого периода будет оправдан более высоким качеством их ранжирования и последующего отбора. На рис. 5 изображены графики, аналогичные таковым на рис. 3 и 4, однако построенные для более длительного периода наблюдения — 5 лет. Здесь, напротив, уже появляется возможность сравнительно достоверно отделить высокопродуктивных ученых от низкопродуктивных. При высокой доле времени, посвящаемой науке, от 60 % и выше, вероятности ошибок первого и второго рода составят:

- для 30%-го порога (вариант В) — соответственно не более 0,3 и 0,35;

- для 70%-го порога (вариант Б) — соответственно не более 0,2 и 0,05.

Наиболее жесткий, 100%-й порог (вариант А), преодолевается учеными-лидерами с вероятностью не более 0,4–0,45. Таким образом, в данном случае наиболее целесообразно устанавливать порог ранжирования на уровне 70 % ожидаемой результативности ученых-лидеров.

Однако при малой доле времени, посвящаемой науке, около 20–30 %, вероятности ошибок первого и второго рода составят:

- для 30%-го порога (вариант В) — соответственно до 0,2 и 0,3–0,45;

- для 70%-го порога (вариант Б) — соответственно 0,35–0,4 и 0,07–0,08.

Следовательно, и при относительно длительном периоде наблюдения низкая обеспеченность ученых и соответственно низкая доля времени, посвящаемая науке, значительно ухудшают качество оценки продуктивности и отбора ученых.

Таким образом, важным фактором — и в предлагаемой здесь модели, и в реальном управлении развитием научного потенциала — является длительность периода материальной поддержки и соответственно длительность наблюдения за результативностью ученых, по итогам которого и будет проводиться их отбор. «Длительный» он или «короткий» — определяется характерными значениями продуктивности ученых. Фактически, необходимо, чтобы ожидаемое число результатов за период наблюдения, по крайней мере, для высокопродуктивных исследователей, существенно превышало единицу. В противном случае неизбежны высокие вероятности ошибок первого и второго рода, что обесмысливает отбор ученых на основании результативности, измеренной за период наблюдения. При установлении длительности периода наблюдения следует учитывать, что характерные значения продуктивности ученых (в том числе уровни публикационной активности) сильно варьируются в зависимости от отрасли науки и даже конкретной ее области. Это касается и характерной длительности цикла исследований.

Вопрос выбора длительности периода наблюдения можно трактовать шире — до каких пор ученых следует считать «молодыми»? В различных грантовых фондах и в статистической отчетности научных и образовательных организаций «порог молодости» для ученых варьируется. Например, в российской государственной статистике и в разнообразных нормативах для государственных вузов и научных организаций весьма распространен порог, равный 39-ти годам, в научных фондах и в положениях о грантах Президента РФ — 40–45 лет для докторов наук и 35 лет — для кандидатов. Учитывая возможность начала научной работы со старших курсов вуза, т. е. приблизительно с 20 лет, это означает, что ученый считается молодым около 15–20 лет с начала карьеры. В реальности на протяжении жизненного цикла истинная продуктивность ученых, как правило, растет благодаря накоплению опыта и росту квалификации. Поэтому за весьма длительный период, пока ученый считается молодым, его продуктивность может (и должна) возрасти многократно благодаря эффекту обучения. В некотором смысле этот эффект сглаживает последствия неправильного «отбора в ученые» — действительно, «дорогу осилит идущий», и вполне возможно, что изначально менее способный работник, оставшись в науке, многократно перерастет в квалификации своих изначально более способных, но отсеявшихся по различным причинам конкурентов. Однако в рамках простейших изложенных здесь моделей можно полагать, что, если отбирать более способных претендентов, то и дальнейшее их «обучение» принесет гораздо больший результат.

ЗАКЛЮЧЕНИЕ

- Одним из возможных объяснений положительного влияния молодежных грантов и других форм финансовой поддержки молодых ученых на развитие науки заключается в том, что такая поддержка позволяет молодым ученым тратить меньше времени на подработку и больше — собственно на исследования. Иначе подработка для молодых исследователей, еще не обладающих высокой квалификацией и статусом, практически неизбежна. При малой доле времени, уделяемой науке, оценки результативности ученых подвержены сильным случайным изменениям, а их последующее ранжирование на высоко- и низкопродуктивных недостоверно, что снижает среднюю продуктивность ученых, прошедших отбор, и среднюю эффективность науки. Наличие финансовой поддержки молодых ученых позволяет им тратить больше времени не на подработку, а собственно на занятие наукой, что повышает ожидаемую результативность их работы, достоверность выделения более и менее продуктивных ученых из общей массы и, как следствие, среднюю эффективность научного сообщества после ранжирования и отбора ученых.
- Количественные оценки показывают, что если за период наблюдения высокопродуктивный ученый может получить около 10 условных результатов при 100 %-й занятости наукой, но ученые тратят на научные исследования около 20–30 % времени, вероятности ошибок первого рода (отсева высокопродуктивных ученых) при ранжировании составят 0,25–0,5 даже при относительно мягких порогах ранжирования (на уровне 30–70 % ожидаемой результативности лидеров за период). Вероятности ошибок второго рода (т. е. успешного прохождения отбора низкопродуктивными учеными) составят 0,07–0,25. В этих условиях наиболее целесообразно устанавливать «мягкий» порог ранжирования на уровне около 30 % ожидаемой результативности лидеров за период. При этом само ранжирование позволит повысить среднюю продуктивность ученых не более чем на 20 %. Если же финансовая поддержка позволяет ученым заниматься наукой 50–60 % времени и более, более целесообразно устанавливать порог ранжирования на уровне 70 % ожидаемой результативности лидеров за период. При этом вероятности ошибок первого и второго рода сократятся, соответственно, до 0,1–0,37 и 0,05–0,28, что позволит повысить среднюю продуктивность ученых по итогам ранжирования на 30–35 %. Оценки получены в оптимистическом предположении, что финансовые возможности и потребности всех ученых одинаковы и различие интенсивности появления результатов обусловлено только различием истинной продуктивности. Реальное различие финансового положения ученых и соответственно доли

времени, которую они могут уделять науке, усугубляет проблему ошибок оценки продуктивности ученых и их последующего ранжирования.

- Увеличение периода наблюдения повышает точность оценки продуктивности ученых, достоверности их ранжирования и отбора. Так, в рассмотренных примерах при увеличении длительности периода наблюдения с 3 до 5 лет вероятности ошибок первого и второго рода сократятся, соответственно, до 0,03–0,2 и 0,05–0,35 в случае, если ученые посвящают науке 50–60 % рабочего времени. Становятся более оправданными более жесткие пороги ранжирования, приближающиеся к ожидаемой результативности лидеров за период наблюдения. В целом, продолжительность периода наблюдения за результативностью ученых (и соответственно продолжительность периода поддержки) должна обеспечивать возможность получения значительного числа условных «результатов», около 10 и более. В противном случае неизбежны высокие вероятности ошибок оценки продуктивности ученых, их ранжирования и отбора.

ЛИТЕРАТУРА

1. *Ильина И.Е.* Новая модель грантового финансирования российской науки // Инновационная экономика и промышленная политика региона (ЭКОПРОМ-2014): тр. междунар. науч.-практ. конф. 15–23 сент. 2014 г. — СПб., 2014. — С. 39–45.
2. *Ильина И.Е., Сергеева О.Л., Юревич М.А.* Механизмы повышения эффективности системы научных фондов в России (коллективная монография) // Экономика и промышленная политика: теория и инструментарий / под ред. д-ра экон. наук, проф. А.В. Бабкина. — СПб.: Изд-во СПб-ГПУ, 2014. С. 32–68.
3. *Семенов Е.В.* Гранты в российской науке: достоинства, недостатки, перспективы // Наукоедение. — 2002. — № 4. — С. 9–10.
4. Игра в цифры, или как теперь оценивают труд ученого: (сб. ст. о библиометрике). — М.: МЦНМО, 2011. — 72 с. — URL: www.mcsme.ru/free-books/bibliometric.pdf (дата обращения: 10.03.2018).
5. *Клочков В.В., Крупина С.М.* Экономический анализ эффективности ранжирования научных работников по наукометрическим критериям // Экономический анализ: теория и практика. — 2013. — № 44 (347). — С. 14–29.
6. *Наукометрия и экспертиза в управлении наукой* // Управление большими системами. — 2013. — Вып. 44. — 568 с. (спец. вып.). — URL: http://ubs.mtas.ru/archive/search_results_new.php?publication_id=19079 (дата обращения: 10.03.2018).
7. *Клочков В.В., Рождественская С.М.* Анализ влияния финансирования научно-исследовательских работ на их эффективность: временные аспекты // Экономический анализ: теория и практика. — 2014. — № 30 (381). — С. 37–50.
8. *Баркалов С.А., Новиков Д.А., Попов С.С.* Индивидуальные стратегии предложения труда: теория и практика. — М.: ИПУ РАН, 2002. — 110 с.
9. *Новиков Д.А.* Экспериментальное исследование индивидуальных стратегий предложения труда. — М.: Эгвес, 2010. — 104 с.

Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым.

Клочков Владислав Валерьевич — д-р экон. наук, канд. техн. наук, вед. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ vlad_klochkov@mail.ru,
Рождественская Софья Михайловна — нач. отдела, Институт им. Н.Е. Жуковского, г. Жуковский, ✉ rozhddestvenskayasm@nrczh.ru.

ПРОТИВОЗАТРАТНЫЕ МЕХАНИЗМЫ ФИНАНСИРОВАНИЯ

А.В. Щепкин

Рассмотрена задача определения объемов финансирования нескольких работ, составляющих общее производственное задание. Исполнитель каждой работы — монополист в соответствующей области и не имеет конкурентов. Фонд финансирования определен. Исследованы механизмы, обеспечивающие получение достоверной информации о затратах на выполнение каждой работы.

Ключевые слова: затраты, прибыль, рентабельность, авансовый платеж.

ВВЕДЕНИЕ

Под противозатратным механизмом здесь понимается такой механизм, применение которого приводит к тому, что исполнителям не выгодно завышать планируемые затраты на выполнение работ. Идея противозатратности и соответственно описания некоторых противозатратных механизмов представлены в работах [1–3]. В централизованной экономике основное требование к предприятиям заключалось в выполнении плана, причем показатели, характеризующие план, далеко не всегда стимулировали снижение затрат. Поэтому разработка противозатратных механизмов, обеспечивающих снижение затрат, были особенно востребованы для подобных ситуаций. Что касается рыночной экономики, то надежды, что стремление предприятий к максимизации своей прибыли путем снижения затрат и «невидимая рука рынка» все расставит по своим местам, в полной мере не оправдались. Подтверждение тому большое количество публикаций об управлении затратами в странах с развитой рыночной экономикой [4–6]. Заметим здесь, что необходимость применения противозатратных механизмов в рыночной экономике может быть вызвано наличием производителей-монополистов и стремлением государства ограничить рост цен на продукцию производителей-монополистов путем введения ограничений на их рентабельность [7]. Кроме того, монопольные эффекты как в рыночной, так и в централизованной экономике, могут проявляться в системе «заказчик — исполнители».

Так, например, монопольный эффект в системе «заказчик — исполнители» может проявляться, ес-

ли исполнители представляют собой узкоспециализированные производства, деятельность которых может только дополнять друг друга, а не дублировать.

Рассмотрим ситуацию, когда заказчику необходимо выполнить некоторое производственное задание и на его выполнение в распоряжении заказчика имеется определенный объем финансирования. Для выполнения задания заказчик заключает договор с исполнителями. При заключении договора заказчику необходимо распределять и работы, и финансирование среди исполнителей. В связи с тем, что исполнители представляют собой узкоспециализированные производства, принципиальных сложностей при распределении работ между ними не возникает. В то же время при распределении финансовых средств каждый исполнитель «тянет одеяло на себя» — стремится получить максимально возможное финансирование, которое позволит ему увеличить свою прибыль. Стремление исполнителей увеличить свою прибыль и, как следствие, желание получить как можно больше финансовых средств может привести к искажению информации, используемой заказчиком при распределении финансовых средств.

1. МОДЕЛЬ ОБЪЕКТА

Модель взаимодействия заказчика и исполнителей представляет собой двухуровневую активную систему [8], состоящую из Центра (заказчика) и множества $N = \{1, 2, \dots, n\}$ агентов нижнего уровня (исполнителей). Центр, на основе выбранного им механизма распределения, определяет размер финансовых средств, который выделяется агентам

на выполнение работ. Задача агентов состоит в выполнении соответствующих работ и получении максимальной прибыли. В настоящей работе прибыль агента определяется как разность между полученным финансированием и фактическими затратами на выполнение работы. Если считать, что на выполнение конкретной работы требуются вполне конкретные затраты, то отсюда следует, чем выше финансирование агента, тем выше его прибыль.

При распределении имеющихся средств Центр запрашивает у агентов информацию о планируемых затратах на выполнение работ. Стратегия агента состоит в сообщении Центру запрашиваемой информации о затратах.

Обозначим: C — фонд финансирования; s_i — планируемые затраты i -го агента; c_i — объем финансирования i -го агента; z_i — фактические затраты i -го агента; $i \in N$.

Плановая прибыль i -го агента определяется выражением $P_{in} = c_i - s_i$, соответственно фактическая прибыль записывается как $P_{if} = c_i - z_i$. Последнее выражение представим в виде $P_{if} = c_i - s_i + s_i - z_i = P_{in} + s_i - z_i$. Разницу $P_{ic} = s_i - z_i$ будем называть сверхплановой прибылью. Здесь будем считать, что планируемые затраты всегда больше фактических затрат, т. е. выполняется неравенство, $s_i \geq z_i$, поэтому $P_{ic} \geq 0$. Кроме того, будем рассматривать случай, когда размер фонда финансирования превышает сумму фактических затрат всех агентов, т. е. справедливо

$$C > \sum_{i \in N} z_i. \quad (1)$$

Также полагаем, что Центр оставляет в распоряжении агента только часть сверхплановой прибыли. При этом выражение для фактической прибыли i -го агента имеет вид $P_{if} = P_{in} + qP_{ic} = c_i - s_i + q(s_i - z_i)$, $i \in N$, где $q \in [0; 1]$ — норматив, определяющий размер сверхплановой прибыли, остающейся у агента.

Будем считать, что целевой функцией Центра является разность между планируемыми и фактическими затратами $\varphi = \sum_{i \in N} (s_i - z_i)$ и Центр, используя различные механизмы распределения фонда финансирования, стремится минимизировать эту разность. Из выражения для целевой функции Центра легко видеть, что она принимает минимальные значения, если Центром будут созданы условия неманипулируемости для агентов [9, 10].

2. ПРИНЦИП РАВНЫХ РЕНТАБЕЛЬНОСТЕЙ

Этот принцип реализует идею выделения каждому агенту такого количества средств, которое обеспечивает агентам одинаковую рентабельность, т. е. получение одинаковой прибыли на единицу вложенных средств [3, 11]. При определении объемов финансирования агентов на основе этого принципа Центр решает задачу

$$\begin{cases} \frac{c_i - z_i}{z_i} = \frac{c_j - z_j}{z_j}, \\ \sum_{i \in N} c_i = C. \end{cases}$$

Решение этой задачи записывается в виде

$$c_i^{(z)} = (\rho^{(z)} + 1)z_i, \quad i \in N, \quad (2)$$

где $\rho^{(z)}$ — рентабельность выполнения работ, определяемая выражением

$$\rho^{(z)} = C / \sum_{j \in N} z_j - 1. \quad (3)$$

Учитывая неравенство (1), можно утверждать, что $\rho^{(z)} > 0$.

Случай 1. Так как Центру не известны точные значения фактических затрат z_i , при определении объемов финансирования он использует информацию о планируемых затратах s_i , $i \in N$, полученную от агентов. Объем финансирования i -го агента определяется как

$$c_i^{(s)} = (\rho^{(s)} + 1)s_i, \quad i \in N, \quad (4)$$

где $\rho^{(s)}$ — плановая рентабельность агентов, определяемая выражением

$$\rho^{(s)} = C / \sum_{j \in N} s_j - 1. \quad (5)$$

Так как $s_i \geq z_i$, то, сравнивая выражения (3) и (5), можно заключить, что всегда справедливо $\rho^{(z)} \geq \rho^{(s)}$. Таким образом, максимальная рентабельность $\rho^{(z)}$ определяется выражением (3).

Утверждение 1. Если $c_i^{(s)}$, $i \in N$, определяются в соответствии с выражением (4) и все агенты сообщили достоверную информацию, т. е. $s_i = z_i$, $i \in N$, то k -му агенту не выгодно увеличивать планируемые затраты, если справедливо неравенство

$$z_k \geq \left(1 - \frac{1-q}{\rho^{(z)} + 1}\right) \frac{C}{\rho^{(z)} + 1}. \quad (6)$$

Доказательство. Так как $c_i, i \in N$, определяются в соответствии с выражением (4), фактическая прибыль i -го агента записывается в виде

$$P_{i\Phi} = \rho^{(s)} s_i + q(s_i - z_i), \quad i \in N. \quad (7)$$

Когда $s_i = z_i, i \in N$, прибыль агентов равна $P_{i\Phi} = \rho^{(z)} z_i, i \in N$. Пусть $s_i = z_i, i = 1, \dots, k-1, k+1, \dots, n$, а $s_k = z_k + \delta$, где $\delta > 0$. Прибыль k -го агента в этом случае составит $P'_{k\Phi} = \rho^{(\delta)}(z_k + \delta) + q\delta$. Нетрудно показать, что в этом случае

$$\rho^{(\delta)} = \frac{C(\rho^{(z)} + 1)}{C + \delta(\rho^{(z)} + 1)} - 1. \quad (8)$$

Покажем, что прибыль k -го агента уменьшилась, т. е. выполняется неравенство $P_{k\Phi} \geq P'_{k\Phi}$. Перепишем это неравенство в виде

$$\rho^{(z)} z_k \geq \rho^{(\delta)}(z_k + \delta) + q\delta. \quad (9)$$

Учитывая зависимость (8), неравенство (9) можем переписать в виде

$$\delta \geq C \frac{\rho^{(z)} + q}{(1-q)(\rho^{(z)} + 1)} - z_k \frac{\rho^{(z)} + 1}{1-q}, \quad (10)$$

так как $\delta \geq 0$, то неравенство (10) будет выполняться всегда, если $C \frac{\rho^{(z)} + q}{(1-q)(\rho^{(z)} + 1)} - z_k \frac{\rho^{(z)} + 1}{1-q} \leq 0$.

Отсюда следует, что неравенство (9) справедливо, если справедливо неравенство (6). Утверждение доказано.

Следствие. Ситуация $s_i^* = z_i, i \in N$, является ситуацией равновесия по Нэшу, если выполняется неравенство

$$\min_i \{z_i\} \geq \left(1 - \frac{1-q}{\rho^{(z)} + 1}\right) \frac{C}{\rho^{(z)} + 1}. \quad \blacklozenge \quad (11)$$

Действительно, если выполняется неравенство (11), то неравенство (6) выполняется для любого $k \in N$, это означает, что ни одному агенту не выгодно увеличивать планируемые затраты.

Таким образом, целевая функция Центра принимает минимальное значение, если справедливо неравенство (11).

Ограничение (11) накладывает достаточно жесткие требования, обеспечивающие существование ситуации равновесия по Нэшу. Перепишем его в виде

$$\min_i \{z_i\} \geq A \left(1 - \frac{1-q}{\rho^{(z)} + 1}\right) n, \quad (12)$$

где $A = \frac{1}{n} \sum_{j \in N} z_j$ — среднее арифметическое значение фактических затрат всех агентов.

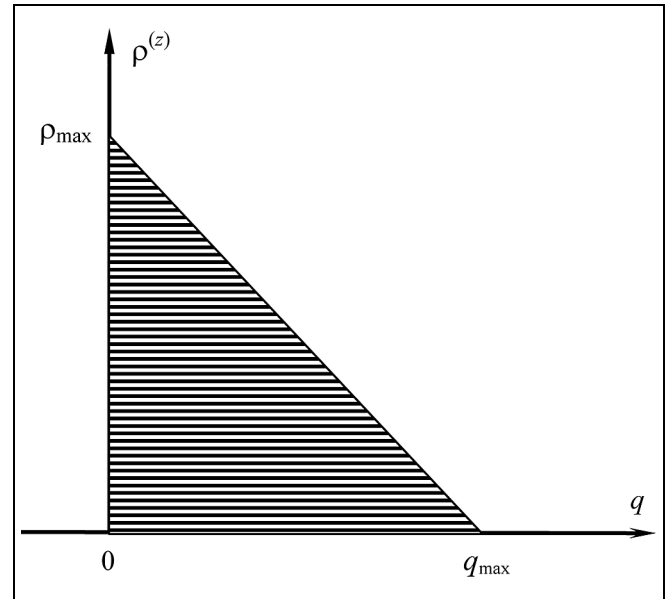


Рис. 1. Область противозатратности

В работе [12] показано, что для любого среднего, в том числе и среднего арифметического значения, справедливо неравенство $\min_{i \in N} \{z_i\} \leq A$.

Из неравенства (12) видно, что среднее арифметическое значение фактических затрат агентов умножается на некоторый коэффициент, такой, что минимальное значение фактических затрат оказывается больше полученного произведения.

Так как в общем случае фактические затраты агентов на выполнение работ разные, поскольку коллектив агентов неоднородный, то будем считать, что степень неоднородности коллектива агентов ω определяется как

$$\omega = \min_i z_i / A. \quad (13)$$

Очевидно, что если фактические затраты агентов одинаковые, то $\omega = 1$, в противном случае всегда справедливо неравенство $\omega < 1$, при этом чем более неоднородный коллектив, тем меньшее значение принимает величина ω .

Учитывая определение (13), неравенство (12) можно переписать в виде

$$\rho^{(z)} \leq (\omega - nq)/(n - \omega). \quad (14)$$

Отсюда следует, что максимальное значение $\rho_{\max}$, которое может принять величина $\rho^{(z)}$, равно $\omega/(n - \omega)$, а максимальное значение $q_{\max}$, которое может принять величина q , равно ω/n . Значения $\rho^{(z)}$ и q , при которых неравенство (14), справедливо находятся в заштрихованной области, изображен-

ной на рис. 1. В дальнейшем эту область будем называть областью противозатратности, а случай 1 — базовым.

Из рис. 1. следует, что чем больше степень неоднородности агентов, т. е. чем меньше значение ω , тем меньше область противозатратности.

3. ИСПОЛЬЗОВАНИЕ АВАНСОВЫХ ПЛАТЕЖЕЙ ДЛЯ ОБЕСПЕЧЕНИЯ ПРОТИВОЗАТРАТНОСТИ

Рассмотрим случай, когда Центр определяет планируемый объем финансирования в соответствии с выражением (4), но сначала выплачивает только аванс, а окончательный расчет осуществляет после завершения работ.

Пусть размер аванса a_i i -го агента составляет p -ую часть объема финансирования $c_i^{(s)}$, т. е. $a_i = pc_i^{(s)}$, где $p < 1$.

Как было отмечено выше, для выполнения работы i -му агенту, $i \in N$, необходимо затратить средства в размере z_i . Если $z_i \geq a_i$, агент для выполнения работы доплачивает свои собственные средства. Если этих средств недостаточно — агент берет кредит в банке.

Случай 2. Пусть собственных средств агента достаточно для выполнения работ. Нетрудно показать, что в этом случае случай 2 аналогичен базовому случаю. Действительно, получив аванс a_i и добавив собственных средств на сумму $z_i - a_i$, i -й агент, $i \in N$, выполняет весь объем работ и затем получает оставшуюся часть финансирования в размере $c_i^{(s)} - a_i$. При этом объем финансирования не меняется и определяется выражением (4). Прибыль i -го агента определяется выражением (7); таким образом, остается справедливым утверждение 1.

Случай 3. Собственными средствами агенты не располагают, для выполнения работ они берут кредит в банке по ставке кредитования, равной β , и поэтому общие затраты агентов на выполнение работы складываются из самих затрат z_i и платы за пользование кредитом $\beta(z_i - a_i)$. В этом случае, общие затраты i -го агента

$$z_{oi} = z_i + \beta(z_i - a_i) = (1 + \beta) z_i - \beta p(\rho^{(s)} + 1)s_i, \quad i \in N. \quad (15)$$

После выполнения работ i -й агент получает оставшуюся часть финансирования в размере $c_i^{(s)} - a_i$, и его прибыль

$$P_{i\Phi} = c_i^{(s)} - s_i + q(s_i - z_i) - \beta(z_i - a_i), \quad i \in N.$$

Учитывая выражения (4), (15) и $a_i = pc_i^{(s)}$, выражение для фактической прибыли можно записать в виде

$$P_{i\Phi} = [(1 + \beta)\rho^{(s)} + q + \beta p]s_i - (q + \beta)z_i, \quad i \in N.$$

Утверждение 1 для этого случая может быть сформулировано как

Утверждение 2. Если $c_i^{(s)}$, $i \in N$, определяются в соответствии с выражением (4), агенты получают сначала аванс, а недостающие средства для выполнения работ покрывают кредитом, и все агенты сообщают достоверную информацию, т. е. $s_i = z_i$, $i \in N$, то k -му агенту не выгодно увеличивать планируемые затраты, если справедливо неравенство

$$z_k \geq \left(1 - \frac{1 - q}{(1 + \beta p)(\rho^{(z)} + 1)}\right) \frac{C}{\rho^{(z)} + 1}. \quad \blacklozenge$$

Утверждение 2 доказывается аналогично доказательству утверждения 1.

Здесь также нетрудно показать, что $s_i^* = z_i$, $i \in N$, является ситуацией равновесия по Нэшу, если выполняется неравенство

$$\min_{i \in N} z_i \geq A \left(1 - \frac{1 - q}{(1 + \beta p)(\rho^{(z)} + 1)}\right) n,$$

из которого, в свою очередь, следует

$$\rho^{(z)} \leq \frac{\omega}{n - \omega} - \frac{n}{n - \omega} \frac{q + \beta p}{1 + \beta p}.$$

Область противозатратности для случая 3 имеет такой же вид, как и на рис. 1. Но в этом случае

$$\rho_{\max} = \frac{\omega}{n - \omega} - \frac{n}{n - \omega} \frac{\beta p}{1 + \beta p}, \quad \text{а } q_{\max} = \frac{\omega}{n} - \frac{n - \omega}{n} \beta p.$$

А это означает, что область противозатратности в случае 3 сократилась по сравнению с базовым случаем.

В существующей практике окончательный расчет производится, когда все работы завершены, и Центр может проверить, какие фактические затраты были у агента на выполнение работ. А это означает, что перед выплатой оставшихся средств Центр, зная фактические затраты, может пересчитать объем финансирования.

Случай 4. Пересчет объемов финансирования осуществляется в соответствии с выражением (2), что соответствует определению объемов финансирования агентов на основе достоверной информации об их затратах.

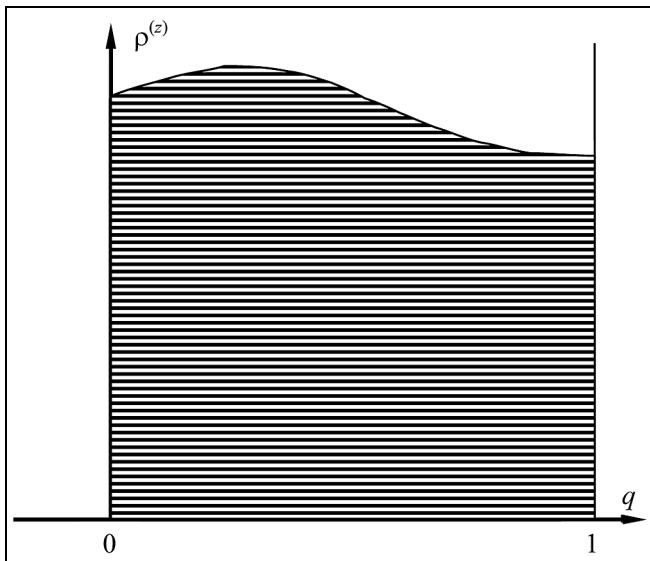


Рис. 2. Область противозатратности при определении объема финансирования в соответствии с выражением (2) и наличия у агентов собственных средств

Если i -й агент использует собственные средства на выполнение работ, его фактическая прибыль определяется как $P_{i\phi} = c_i^{(z)} - s_i + q(s_i - z_i)$.

Учитывая выражения (2) и (3), прибыль агента можно записать в виде $P_{i\phi} = (\rho^{(z)} + 1 - q)z_i - (1 - q)s_i$.

Отсюда легко видеть, что для получения максимальной прибыли агентам не выгодно увеличивать планируемые затраты при любом значении $\rho^{(z)}$ и при любом значении $q \in [0; 1)$, и, соответственно, в ситуации равновесия по Нэшу $s_i^* = z_i$, $i \in N$.

Область противозатратности для этого случая имеет вид заштрихованной фигуры, показанной на рис. 2.

Случай 5. Если у агентов собственные средства на выполнение работ отсутствуют, они используют заемные средства. Затраты i -го агента за пользование кредитом будут равны $\beta(z_i - a_i)$, и его прибыль будет определяться как

$$P_{i\phi} = c_i^{(z)} - s_i + q(s_i - z_i) - \beta(z_i - pc_i^{(s)}), \quad i \in N.$$

Подставляя в это выражение значения $c_i^{(z)}$ и $c_i^{(s)}$, получаем

$$P_{i\phi} = (\rho^{(z)} + 1 - q - \beta)z_i - [1 - q - \beta(\rho^{(s)} + 1)]s_i.$$

Утверждение 1 для этого случая может быть сформулировано как

Утверждение 3. В случае, когда агенты получают аванс, равный p -й части $c_i^{(s)}$, и для выполнения работ берут кредит, а окончательный объем финансирования определяется в соответствии с выражением (2), то при сообщении всеми агентами достоверной информации, т. е. $s_i = z_i$, $i \in N$, k -му агенту не выгодно увеличивать планируемые затраты, если справедливо неравенство

$$z_k \geq \left[1 - \frac{1 - q}{\beta p(\rho^{(z)} + 1)} \right] \frac{C}{\rho^{(z)} + 1}. \quad \blacklozenge$$

Утверждение 3 доказывается аналогично доказательству утверждения 1.

Здесь также нетрудно показать, что ситуация $s_i^* = z_i$, $i \in N$, является ситуацией равновесия по Нэшу, если выполняется неравенство

$$\min_{i \in N} z_i \geq A \left[1 - \frac{1 - q}{\beta p(\rho^{(z)} + 1)} \right] n.$$

Из этого неравенства нетрудно получить

$$\rho^{(z)} \leq \frac{\omega}{n - \omega} + \frac{n}{n - \omega} \frac{1 - q - \beta p}{\beta p}.$$

Область противозатратности для случая 5 имеет такой же вид, как показано на рис. 1. В этом случае

$$\rho_{\max} = \frac{\omega}{n - \omega} + \frac{n}{n - \omega} \frac{1 - \beta p}{\beta p}, \quad \text{а } q_{\max} = \frac{\omega}{n} + \frac{n - \omega}{n} (1 - \beta p).$$

А это означает, что область противозатратности в случае 5 увеличилась по сравнению с базовым случаем. Причем чем меньше аванс и меньше ставка кредита, тем больше область противозатратности.

Другими словами, при финансировании агентов на основе фактических затрат, даже если к затратам на выполнение работ добавляются затраты за пользование кредитом $\beta(z_i - a_i)$, требования к существованию ситуации равновесия по Нэшу, в которой агенты сообщают достоверную информацию о своих затратах, становятся менее ограниченными.

В рассмотренных выше случаях равные рентабельности агентам определяются только на этапе планирования, и агенты имеют одинаковые рентабельности на этапе реализации, только когда их планируемые затраты совпадают с фактическими, т. е. в области противозатратности.

Пусть теперь на этапе планирования Центр в первую очередь фиксирует рентабельности агентов и определяет размер их аванса. Фиксация рентабельности здесь имеет принципиальное значение, так как именно эта зафиксированная рентабель-

ность используется для окончательного расчета за выполнение работ.

Случай 6. Итак, на этапе планирования определяется размер аванса i -го агента $a_i = pc_i^{(s)}$ и определяется рентабельность выполнения работ $\rho^{(s)}$ в соответствии с выражением (5).

При окончательном расчете, когда фактические затраты z_i уже известны, финансирование агентов определяется из условия

$$c_i = (\rho^{(s)} + 1)z_i, \quad i \in N. \quad (16)$$

Очевидно, что размер аванса не должен превышать размер финансовых средств при окончательном расчете, т. е. должно выполняться неравенство $pc_i^{(s)} \leq (\rho^{(s)} + 1)z_i$. А отсюда следует, что $p \leq \min_i \{z_i/s_i\}$.

Если i -й агент использует только собственные средства, его прибыль определяется выражением

$$\begin{aligned} P_{i\phi} &= (1 + \rho^{(s)})z_i - s_i + q(s_i - z_i) = \\ &= (1 + \rho^{(s)} - q)z_i - (1 - q)s_i, \quad i \in N. \end{aligned}$$

Отсюда легко видеть, что максимум своей прибыли для любых значений $\rho^{(z)}$ и $q \in [0; 1)$ агенты получают, когда планируемые затраты равны фактическим затратам, и в ситуации равновесия по Нэшу $s_i^* = z_i, i \in N$.

Область противозатратности для этого случая изображена на рис. 2, из которого следует, что область противозатратности не зависит от значений $\rho^{(z)}$ и q . Это объясняется тем фактом, что если у агента нет необходимости использовать заемные средства, ему не приходится выплачивать проценты по кредиту и его полные затраты на выполнение работы определяются только значением z_i . А из условий (16) легко видеть, что агент получает максимальное финансирование, когда планируемые затраты совпадают с фактическими.

Если же i -й агент использует заемные средства для выполнения работ, то его аванс тем больше, чем больше планируемые затраты s_i , и тем меньше его затраты за пользование кредитом. Соответственно прибыль i -го агента определяется как

$$P_{i\phi} = (1 + \rho^{(s)})z_i - s_i + q(s_i - z_i) - \beta(z_i - pc_i^{(s)}), \quad i \in N.$$

Подставляя в это выражение значение $c_i^{(s)}$, получаем

$$P_{i\phi} = (1 + \rho^{(s)})(z_i + \beta ps_i) - (1 - q)s_i - (q + \beta)z_i, \quad i \in N.$$

Утверждение 1 для этого случая может быть сформулировано как

Утверждение 4. В случае, когда агенты получают аванс, равный p -й части $c_i^{(s)}$, и для выполнения работ берут кредит, а окончательный объем финансирования определяется в соответствии с условием (16), то при сообщении всеми агентами достоверной информации, т. е. $s_i = z_i, i \in N$, k -му агенту не выгодно увеличивать планируемые затраты, если справедливо неравенство

$$z_k \geq \frac{\beta p}{1 + \beta p} \left[1 - \frac{1 - q}{\beta p(\rho^{(z)} + 1)} \right] \frac{C}{\rho^{(z)} + 1} \cdot \diamond$$

Утверждение 4 доказывается аналогично доказательству утверждения 1.

Здесь, как и выше, легко показать, что ситуация $s_i^* = z_i, i \in N$, является ситуацией равновесия по Нэшу, если выполняется неравенство

$$\min_{i \in N} z_i \geq A \frac{\beta p}{1 + \beta p} \left[1 - \frac{1 - q}{\beta p(\rho^{(z)} + 1)} \right] n,$$

из которого получаем

$$\frac{1 - q}{\rho^{(z)} + 1} \geq \frac{(n - \omega)\beta p - \omega}{n}. \quad (17)$$

Отсюда сразу видно, что если $\beta p \leq \omega/(n - \omega)$, то область противозатратности имеет вид, показанный на рис. 2. Чтобы вид области противозатратности не изменялся при уменьшении степени неоднородности коллектива ω , произведение βp при этом также должно уменьшаться.

Если же произведение βp таково, что выполняется условие $\beta p > \omega/(n - \omega)$, то неравенство (17) можно переписать в виде

$$\rho^{(z)} \leq \frac{\omega}{n - \omega} + \frac{n}{n - \omega} \frac{(1 - q - \beta p)(n - \omega) + \omega}{\beta p(n - \omega) - \omega}.$$

Отсюда следует, что область противозатратности имеет вид, показанный на рис. 1. В этом

случае $\rho_{\max} = \frac{\omega}{n - \omega} + \frac{n}{n - \omega} \frac{(1 - \beta p)(n - \omega) + \omega}{\beta p(n - \omega) - \omega}$, а

$$q_{\max} = \frac{\omega}{n} + 1 - \frac{(n - \omega)\beta p}{n}.$$

А это означает, что область противозатратности в случае 6 увеличилась по сравнению с базовым случаем. Причем при уменьшении коэффициента p и уменьшении ставки кредита β увеличивается область противозатратности, т. е. этот случай, в смысле зависимости размера области противозатратности от произведения βp , эквивалентен случаю 5.

4. ПРОТИВОЗАТРАТНОСТЬ НА ОСНОВЕ МАКСИМАЛЬНОЙ И МИНИМАЛЬНОЙ РЕНТАБЕЛЬНОСТЕЙ

При реализации этого механизма Центр определяет и сообщает всем агентам единый минимальный норматив рентабельности ρ_0 , который означает, что каждый агент гарантированно получит такое финансирование от Центра, которое обеспечит ему рентабельность не меньше ρ_0 . Очевидно, что для обеспечения этой гарантии должно выполняться условие $C / \sum_{j \in N} s_j - 1 \geq \rho_0$. На основе полученных от агентов значений планируемых затрат Центр рассчитывает максимальный объем финансирования, который агенты могли бы получить. Максимальный объем финансирования i -го агента, $i \in N$,

$$L_i = C - (1 + \rho_0) \left(\sum_{j \in N} s_j - s_i \right). \quad (18)$$

Фактически это означает, что если бы все агенты, кроме i -го, получали финансирование на основе минимального норматива рентабельности ρ_0 , все оставшиеся средства Центр передавал бы i -му агенту.

Главная особенность максимального объема финансирования i -го агента состоит в том, что он не зависит от значения планируемых затрат самого i -го агента.

Максимальный объем финансирования агента позволяет определить его максимальную рентабельность работы, $\eta_i = (L_i - s_i) / s_i$.

Фактическая рентабельность $\rho_{i\Phi}$ i -го агента, $i \in N$, определяется как линейная комбинация минимальной ρ_0 и максимальной η_i рентабельностями: $\rho_{i\Phi} = (1 - k)\rho_0 + k\eta_i$, $k \in [0, 1]$.

На основе фактической рентабельности в соответствии с выражением (4) определяется объем финансирования i -го агента

$$c_i = (1 + \rho_{i\Phi})s_i = (1 - k)(1 + \rho_0)s_i + kL_i, \quad i \in N. \quad (19)$$

Прибыль i -го агента с учетом выражения (7) может быть записана как

$$P_{i\Phi} = [(1 - k)\rho_0 - k + q]s_i + kL_i - qz_i, \quad i \in N.$$

Отсюда следует, что агентам невыгодно будет завышать планируемые затраты на выполнение работ, если $[(1 - k)\rho_0 - k + q] < 0$. Это неравенство можно переписать в виде

$$\rho_0 < (k - q) / (1 - k). \quad (20)$$

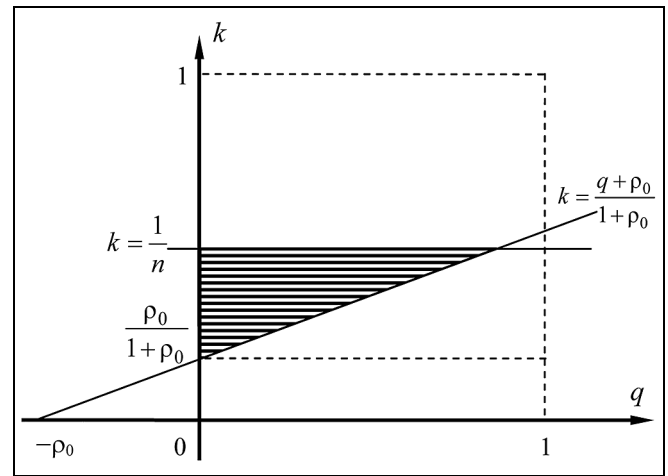


Рис. 3. Область противозатратности для $\rho_0 > 0$

Для реализуемости механизма необходимо, чтобы $\sum_{i \in N} c_i \leq C$. Подставив в это неравенство выражение (19), получим $(1 - kn)(1 + \rho_0) \sum_{i \in N} s_i + knC \geq C$, а отсюда следует $k \leq 1/n$.

Таким образом, при выполнении условий (20) и $k \leq 1/n$ агентам не выгодно будет завышать планируемые затраты, и прибыль i -го агента

$$P_{i\Phi} = \rho_0 z_i + k \left[C - (1 + \rho_0) \sum_{j \in N} z_j \right], \quad i \in N. \quad (21)$$

Заметим, что выбор $\rho_0 > 0$ соответствует тому, что Центр гарантирует i -му агенту получение прибыли в размере $\rho_0 z_i$, $i \in N$, в то время как фактическая прибыль определяется выражением (21) и принимает большее значение, чем гарантированная прибыль. Область противозатратности для этого случая, представлена на рис. 3 в виде заштрихованной области.

Если $\rho_0 = 0$, то Центр гарантирует агентам компенсацию всех запланированных затрат на выполнение работ, в то время как фактическая прибыль $P_{i\Phi} = k \left(C - \sum_{j \in N} z_j \right)$, $i \in N$.

Область противозатратности для этого случая, представлена на рис. 4 в виде заштрихованной области.

Если $-1 < \rho_0 < 0$, то Центр гарантированно компенсирует только часть затрат, в то время как фактическая прибыль определяется выражением (21).

Область противозатратности для этого случая представлена на рис. 5 в виде заштрихованной области.

Несмотря на то, что при $\rho_0 < 0$ Центр гарантирует компенсацию только части затрат, фактический объем финансирования может обеспечить агенту получение прибыли. Действительно, из выражения (21) следует, что i -й агент, $i \in N$, получает прибыль при $\rho_0 < 0$, если справедливо условие $C > (1 + \rho_0) \sum_{j \in N} z_j - \frac{\rho_0}{k} z_i$.

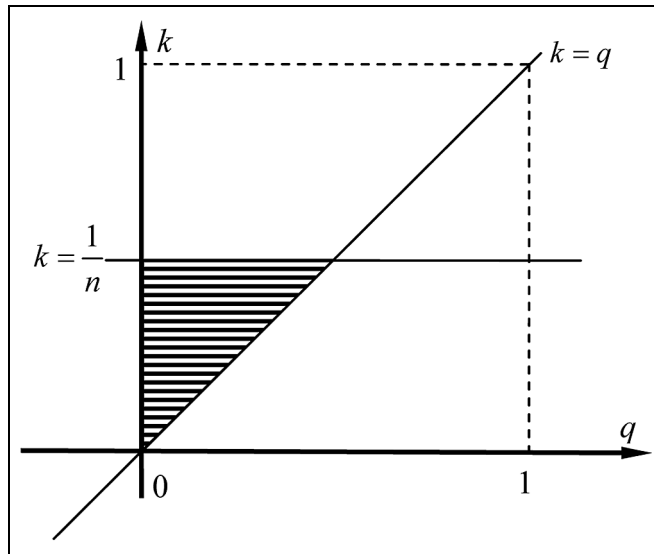


Рис. 4. Область противозатратности для $\rho_0 = 0$

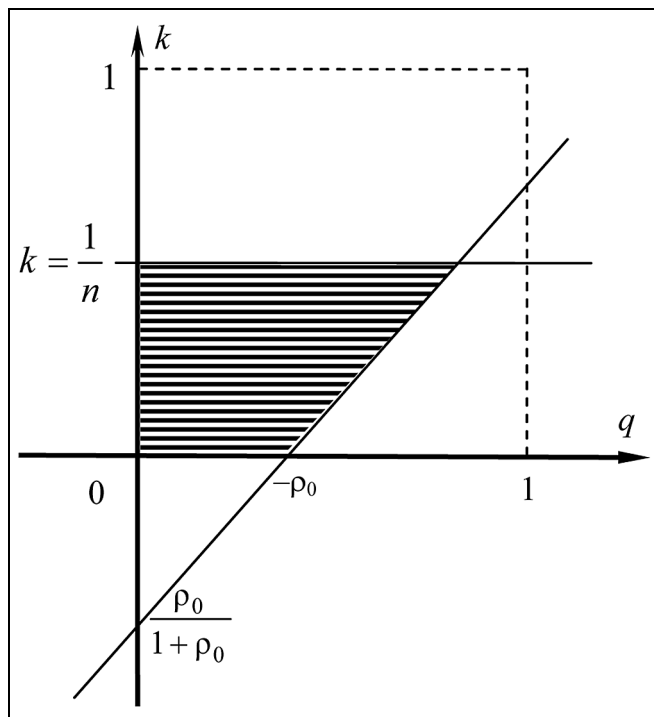


Рис. 5. Область противозатратности для $-1 < \rho_0 < 0$

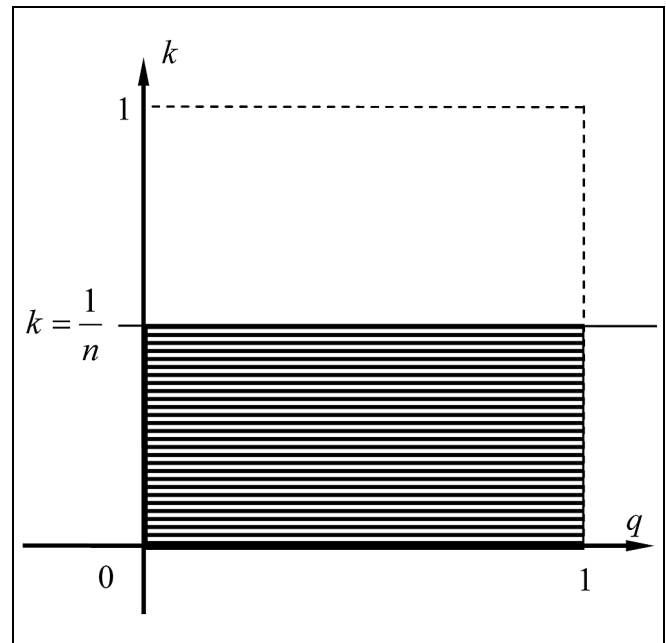


Рис. 6. Область противозатратности для $\rho_0 = -1$

И, наконец, если $\rho_0 = -1$, то это соответствует случаю, когда Центр вообще не гарантирует компенсации затрат агентам на выполнение работ. При этом объем финансирования всех агентов, как следует из выражения (19) с учетом формулы (18), будет составлять $c_i = kC$, $i \in N$, т. е. все агенты будут получать одинаковый объем финансирования, несмотря на то, что фактические затраты на выполнение работ могут значительно отличаться друг от друга. Соответственно, все агенты будут иметь прибыль в размере $P_{i\phi} = kC - z_i$, $i \in N$. Отсюда можно сделать вывод, что все агенты будут работать с прибылью, если справедливо неравенство $\max_i \{z_i\} < kC$.

Область противозатратности для заданного Центром значения $\rho_0 = -1$ представлена на рис. 6 в виде заштрихованной области.

Для построения области противозатратности в осях координат ρ_0 и q перепишем выражение (20) в виде

$$k > (\rho_0 + q)/(1 + \rho_0). \quad (22)$$

Напомним, что $k \leq 1/n$, причем фонд C распределяется между агентами полностью, если $k = 1/n$. Подставив в неравенство (22) значение k , получим

$$\rho_0 < (1 - nq)/(n - 1). \quad (23)$$

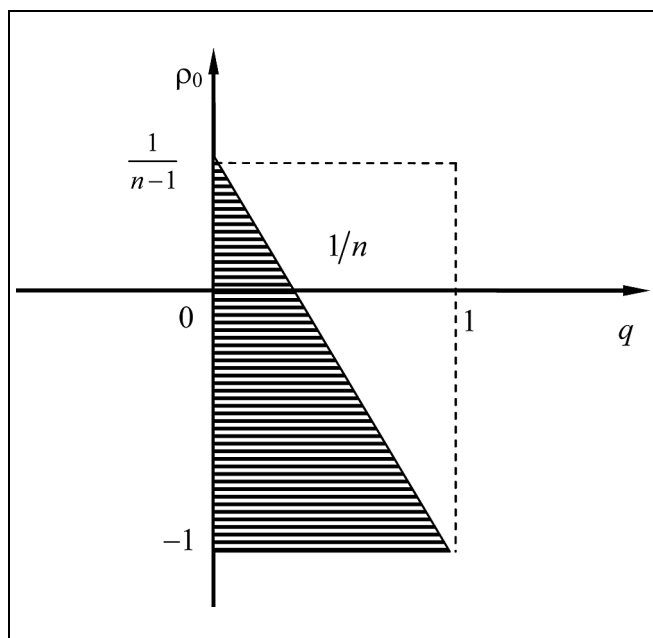


Рис. 7. Область противозатратности

Область противозатратности в осях координат ρ_0 и q представлена на рис. 7 в виде заштрихованной области.

Таким образом, если q и ρ_0 принадлежат заштрихованной области, то агентам выгодно сообщать Центру фактические значения затрат на выполнение работ, и целевая функция Центра будет принимать минимальное значение. Нетрудно также заметить, что с ростом числа агентов n область противозатратности сокращается.

ЗАКЛЮЧЕНИЕ

Определение области противозатратности при распределении финансовых средств на основе принципа равных рентабельностей основывается на прогнозных значениях рентабельности $\rho^{(z)}$ и степени неоднородности агентов ω , так как истинные значения затрат Центру не известны. Поэтому возможность получения достоверной информации о затратах агентов может быть определена таким образом: «Агенты заинтересованы сообщать достоверную информацию о затратах, если значения $\rho^{(z)}$, ω и заданный Центром норматив q удовлетворяют условию (14)». В то же время, при распределении финансовых средств на основе максимальной и минимальной рентабельностей Центр сам формирует область противозатратности, задавая соответствующие параметры, а именно, единый ми-

нимальный норматив рентабельности ρ_0 , значение коэффициента $k \in [0; 1]$ и значение норматива q . Главное, чтобы значения этих параметров удовлетворяли условиям (20) и $k \leq 1/n$.

Отметим, что неравенство (23) — это ограничение на минимальную рентабельность, в то время как при распределении финансовых средств на основе принципа равных рентабельностей выражение (14) — это ограничение на максимальную рентабельность.

ЛИТЕРАТУРА

1. Бурков В.Н., Кашенков А.Р. Противозатратные механизмы управления научными исследованиями и разработками // Совершенствование организационно-экономического механизма управления деятельностью научно-исследовательских и проектно-конструкторских учреждений. — М., 1988. — С. 45.
2. Бурков В.Н., Кашенков А.Р. Противозатратный механизм управления предприятием // Механизмы управления социально-экономическими системами / ИПУ РАН. — 1988. — С. 5—9.
3. Щепкин А.В. Внутрифирменное управление (модели и механизмы): учеб. пособие. — М.: ИПУ РАН, 2001. — 80 с.
4. Anderson S.A. Framework for Assessing Cost Management System Changes: The Case of Activity-Based-Costing Implementation at General Motors // Journal of Management Accounting Research. — 1995. — Vol. 7. — P. 1—51.
5. Mevellec P. Cost Systems Design. — UK: Palgrave Macmillan, 2009. — 311 p.
6. Berk J. Cost Reduction and Optimization for Manufacturing and Industrial Companies. — John Wiley & Sons, Inc. Hoboken, NJ and Scrivener Publishing LLC, Salem, Massachusetts, 2010. — 272 p.
7. Механизмы управления: учеб. пособие / Под ред. Д.А. Новикова. — М.: ЛЕНАНД, 2011. — 192 с.
8. Бурков В.Н. Основы математической теории активных систем. — М.: Наука, 1977. — 256 с.
9. Коргин Н.А. Неманипулируемые механизмы планирования в организационных системах при нетрансферабельной полезности: новые результаты и перспективы // XII Всерос. совещание по проблемам управления ВСПУ-2014 / Москва, ИПУ РАН. — М., 2014. — С. 5325—5331.
10. Бондарик В.Н., Коргин Н.А. Механизмы распределения ресурсов на основе неманипулируемых симметричных анонимных процедур голосования с делегированием // Проблемы управления. — 2012. — № 5. — С. 26—32.
11. Ануфриев И.К., Бурков В.Н., Вилкова Н.И., Рапацкая С.Т. Модели и механизмы внутрифирменного управления. — М.: ИПУ РАН, 1994. — 72 с.
12. Джинни К. Средние величины. — М.: Статистика, 1979. — 447 с.

Статья представлена к публикации руководителем РРС Г.А. Угольником.

Щепкин Александр Васильевич — д-р техн. наук, гл. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, sch@ipu.ru.

БЕСКОНФЛИКТНАЯ САМОМАРШРУТИЗАЦИЯ ДЛЯ ТРЕХМЕРНОГО ОБОБЩЕННОГО ГИПЕРКУБА

В.С. Подлазов

Предложена системная сеть в виде неблокируемого трехмерного обобщенного гиперкуба. Разработана новая структура гиперкуба и для нее предложен алгоритм бесконфликтной прокладки прямых путей между узлами посредством самомаршрутизации в них.

Ключевые слова: системная сеть, обобщенный гиперкуб, полное мультикольцо, коммутационные свойства, неблокируемая сеть, бесконфликтная самомаршрутизация, прямые каналы.

ВВЕДЕНИЕ

Обобщенный r -мерный p -ичный гиперкуб [1] используется как системная сеть в некоторых современных многопроцессорных вычислительных системах [2]. Возможность передачи данных от источников к приемникам по прямым каналам (без буферизации пакетов в промежуточных узлах) повышает быстродействие любой системной сети. Использование прямых каналов необходимо предполагает возможность бесконфликтной маршрутизации пакетов при их произвольной перестановке между узлами.

Сети с бесконфликтной маршрутизацией делятся на два класса — перестраиваемые и неблокируемые. В перестраиваемых сетях бесконфликтная маршрутизация достигается только благодаря составлению соответствующего расписания для каждой перестановки отдельно. Известно, что p -ичный гиперкуб любой размерности r с удвоенным набором межузловых каналов является перестраиваемым для p - p -перестановок (одновременно реализуемых p разных перестановок) [3]. Такой гиперкуб имеет $N = p^r$ узлов и содержит $M = 2Nr(p - 1)$ симплексных каналов. Обобщенный гиперкуб с редукцией части каналов может оставаться перестраиваемым для обычной 1-1-перестановки. В этом случае он содержит $M = 2N(p - 1)[1 + (r - 1)/p]$ симплексных каналов, но становится ассиметричным по числу каналов в разных измерениях.

В неблокируемых сетях имеется возможность прокладки новых бесконфликтных путей, не меняя ранее проложенных путей. Однако для большинс-

тва таких сетей неизвестны алгоритмы бесконфликтной параллельной самомаршрутизации пакетов, при которой каждый пакет прокладывает свой путь через узлы самостоятельно без взаимодействия с пакетами в других узлах. Немногочисленными исключениями служат сети с топологией квазиполных двудольных графов или орграфов, к которым принадлежат двумерные обобщенные гиперкубы и двумерные полные мультикольца [4].

В настоящей работе исследуется возможность построения неблокируемого трехмерного p -ичного гиперкуба и предлагается алгоритм динамической бесконфликтной самомаршрутизации в нем для произвольной перестановки пакетов. Она осуществляется только на основе локальной внутриузловой информации о возникающих конфликтах.

1. НЕБЛОКИРУЕМЫЙ ДВУМЕРНЫЙ ОБОБЩЕННЫЙ ГИПЕРКУБ

Двумерный p -ичный гиперкуб — одна из сетей, которая является как неблокируемой, так и самомаршрутизируемой. В нем каждая строка или столбец, состоящие из p узлов, образуют полный граф. Он имеет $N = p^2$ узлов.

Пример двумерного обобщенного гиперкуба с $N = 9$ узлами приведен на рис. 1. Хотя он выглядит как граф, но его можно представить и как орграф, если учесть, что каждый узел содержит, кроме абонента, еще и коммутатор $p \times p$ (рис. 2), и что абоненты связаны дугами только с коммутаторами [4]. Выходные дуги абонентов направлены только го-

торых есть дуги к приемнику. В примере это коммутаторы 2, 5 и 8. В левой части таблицы (каналы от абонентов) для этих коммутаторов находится тот, к которому есть дуга от источника. По свойствам квазиполных орграфов такой коммутатор один. В примере это коммутатор 5.

По описанной схеме построен коммутатор YARC фирмы CRAY [5]. Он представляет собой СБИС полного коммутатора 64×64 , построенный как двумерный 8-ичный гиперкуб внутри СБИС из 64 коммутаторов 8×8 .

2. ДВУМЕРНЫЕ ГИПЕРКУБЫ В ТРЕХМЕРНОМ ОБОБЩЕННОМ ГИПЕРКУБЕ

Рассмотрим трехмерный p -ичный гиперкуб, содержащий $N = p^3$ узлов, который имеет в каждом узле по $(p - 1)$ выходных и входных дуг для каждого измерения. Эти дуги задают межузловые симплексные каналы.

Разместим трехмерный p -ичный гиперкуб в прямоугольной системе координат так, чтобы в начале координат оказался ближний левый нижний узел. Узлы на оси абсцисс X перенумеруем числами $1, 2, 3, \dots, p$, а узлы на оси ординат Y — числами $1, 1 + p, 1 + 2p, \dots, 1 + (p - 1)p$. Наконец, узлы на оси аппликат Z перенумеруем числами $1, 1 + p^2, 1 + 2p^2, \dots, 1 + (p - 1)p^2$. На рис. 4 приведен пример трехмерного троичного гиперкуба с заданной нумерацией.

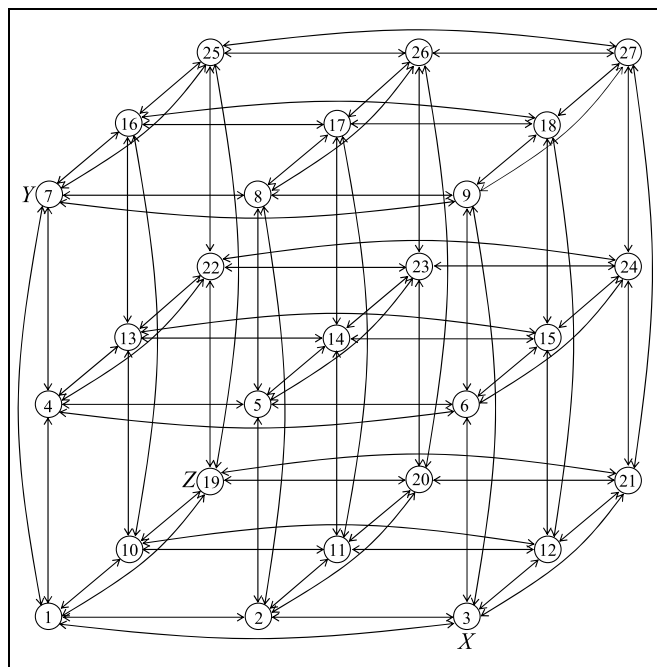


Рис. 4. Трехмерный троичный гиперкуб

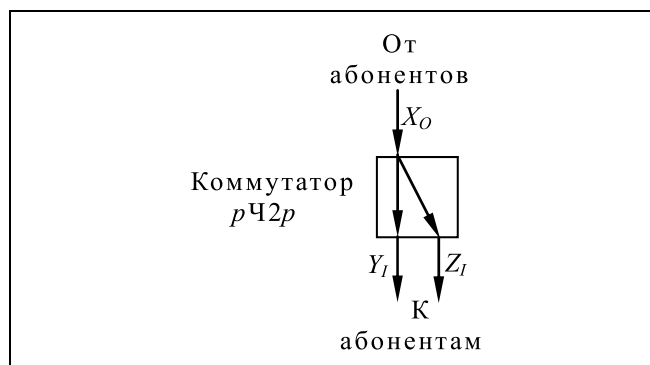


Рис. 5. Коммутатор для двухэтапных прямых путей

Примем как в § 1, что каждый узел содержит абонента и коммутатор с входными/выходными дугами в каждом измерении.

Дуги гиперкуба, которые располагаются «параллельно» оси абсцисс, будем относить к X типу, «параллельно» оси ординат — к Y типу и «параллельно» оси аппликат — к Z типу.

Любой узел гиперкуба принадлежит трем его граням: вертикальной V с дугами X и Y , горизонтальной H с дугами X и Z и боковой L с дугами Y и Z .

Разные грани, содержащие узел a , обозначим как $V(a)$, $H(a)$ и $L(a)$ соответственно. Сам трехмерный гиперкуб можно представить как набор p граней: или $V(1), V(1 + p^2), \dots, V(1 + (p - 1)p^2)$, или $H(1), H(1 + p), \dots, H(1 + (p - 1)p)$, или $L(1), L(2), \dots, L(p)$.

Будем представлять любой путь как последовательность дуг XYZ .

Примем, что любой узел имеет выходные дуги X_0 от абонентов к коммутаторам и входные дуги Y_1 и Z_1 от коммутаторов к абонентам — $3p$ дуг. Из них $3(p - 1)$ дуг являются межузловыми и 3 дуги — локальными внутриузловыми.

Перечисленные дуги делают грани V и H неблокируемыми двумерными гиперкубами. Это означает, что на каждой из них между любыми двумя узлами имеются двухэтапные прямые пути, прокладываемые посредством самомаршрутизации через один коммутатор, представленный на рис. 5. Он имеет схему $p \times 2p$ со сложностью $2p^2$ точек коммутации.

Рассмотренные выше дуги не позволяют представить грань L как неблокируемый двумерный гиперкуб, так как узлы не имеют дуг Y или дуг Z от абонентов к коммутаторам. Поэтому пути между узлами на грани L будем представлять как частный случай трехэтапных путей с локальными дугами на первом или на последнем этапах. Аналогично, пу-

ти между узлами граней V и H , которые не принадлежат им обеим, будем представлять как трехэтапные пути.

Для прокладки трехэтапных путей дополнительно потребуется иметь в каждом узле $p - 1$ основных дуг Y_M , проложенных только между коммутаторами узлов. Между узлами каждого столбца гиперкуба эти дуги составляют полный граф. Вообще, кратчайший путь между любыми двумя абонентами (источником и приемником) можно представить как состоящий из дуги X_O от источника к коммутатору, дуги Y_M между коммутаторами и дуги Z_I от коммутатора к приемнику. При этом двухэтапные и трехэтапные пути вообще не могут иметь конфликтов, так как на первом и на последнем этапах при перестановке используются единственные дуги, а в середине трехэтапных путей — дуги, которые отсутствуют в двухэтапных путях.

Рассмотрим кратчайший трехэтапный путь от узла-источника s к узлу-приемнику d через промежуточные узлы t и i . Первый этап от источника в узле s в промежуточный узел t проходит по дуге X_O . Через узел t могут проходить пути от разных источников, которые на втором этапе используют разные дуги Y_M . Если каждую дугу использует только один путь через узел i , то он является бесконфликтным, и бесконфликтным же он останется на последнем этапе по разным дугам Z_I .

Ситуация, при которой на одну дугу Y_M претендуют несколько путей, считается конфликтом на этой дуге. При этом один из этих путей имеет приоритет и проходит на выход коммутатора, а остальные остаются конфликтными в узле. На дугах Y_M могут возникать конфликты первого типа. Любой такой конфликт возникает, если на прохождение через узел t претендуют несколько путей (рис. 6) из разных узлов-источников s и s^* . Здесь и далее для краткости мы обозначаем только два таких узла, которых на самом деле может быть любое число от 2 до $p - 1$. Пусть эти пути через узел t претендуют на одну и ту же дугу Y_M , ведущей в узел i . На рис. 6 конфликтная дуга обозначена пунктиром, и она при разрешении конфликта не используется.

Для разрешения конфликта в каждом узле необходимы еще $p - 1$ дуг Y_D , которые дублируют дуги Y_M и обозначены на рис. 6 сплошными линиями. Использование дуг Y_D позволяет избежать повторных конфликтов между дугами Y_M и Y_D . Конфликт разрешается таким образом: по заведомо разным дугам Y_D из узла t прокладываются пути в разные вспомогательные узлы h и h^* (узлы без заливки на рис. 6) и осуществляется возврат из них

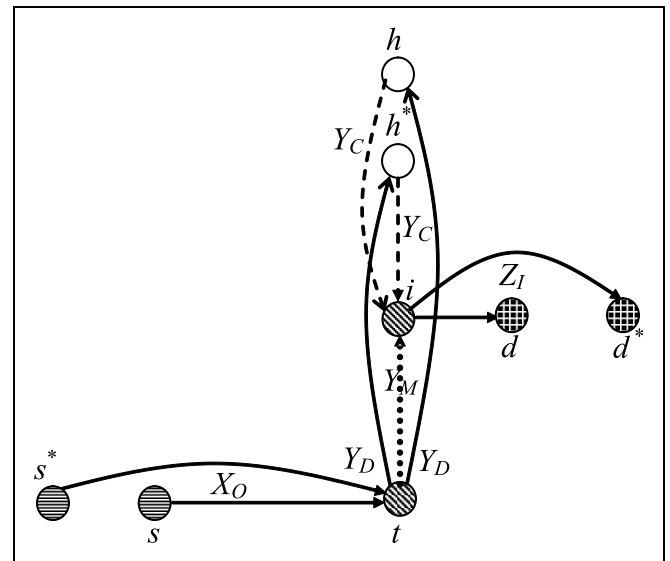


Рис. 6. Конфликт первого типа и его разрешение

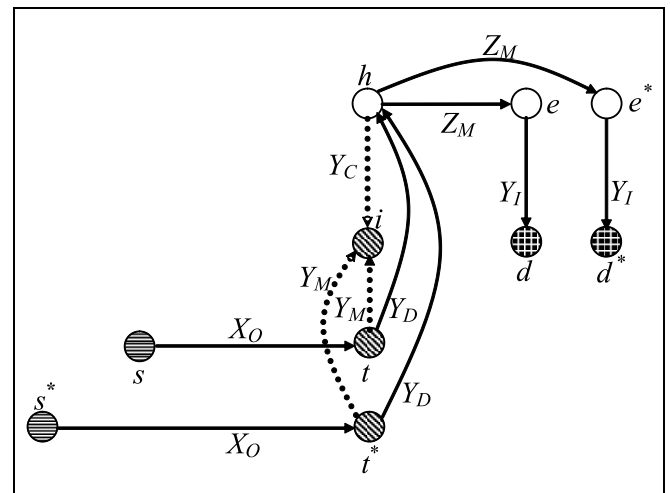


Рис. 7. Конфликт второго типа и его разрешение

в узел i по разным встречным дугам Y_C . Если на дугах Y_C нет конфликтов, то проложенные пути бесконфликтные. Они останутся бесконфликтными и на последнем этапе от коммутаторов узла i к приемникам в узлах d и d^* , так как прокладываются по разным дугам Z_I . Заведомо разные те дуги Y_D , которые являются аналогами разных дуг X_O из узлов s и s^* в узел t .

Однако и на дугах Y_C могут возникать конфликты второго типа. Каждый из них происходит в том случае, если на проход через узел i на втором этапе претендуют несколько путей из разных узлов t и t^* после первого этапа (рис. 7). Для его разрешения

в каждом узле потребуются иметь $p - 1$ дуг Z_M , проложенных только между коммутаторами узлов.

Конфликт второго типа разрешается таким образом: из промежуточного узла h пути прокладывается по тем дугам Z_M , которые равны дугам Z_I последнего этапа. Все эти дуги Z_I различны, поэтому различны и соответствующие дуги Z_M . Поэтому пути бесконфликтно прокладываются в разные вспомогательные узлы e и e^* . Из них маршруты на последнем этапе завершаются бесконфликтно по дугам Y_I из узлов e и e^* .

Конфликты разных типов возникают при разных вариантах размещения узлов-источников на гранях L — на одной грани и на разных гранях. Других вариантов их размещения нет. Поэтому нет и других типов конфликтов.

Рассмотренная процедура задает общую схему прокладки бесконфликтных прямых путей посредством самомаршрутизации, так как осуществляется только на основе локальной информации о конфликтах внутри каждого узла. Наличие такой процедуры самомаршрутизации позволяет сформулировать

Утверждение. Построенный трехмерный p -ичный гиперкуб с $M = 7N(p - 1)$ межузловыми дугами является неблокируемой сетью, для которой предложена процедура бесконфликтной самомаршрутизации. Этот гиперкуб в каждом узле имеет выходные дуги X_O от абонентов, входные дуги Y_I и Z_I к абонентам и дуги Y_M , Y_D , Y_C и Z_M между коммутаторами. ♦

Построенный гиперкуб ассиметричен по дугам разных измерений, и его асимметрия напоминает асимметрию редуцированного перестраиваемого гиперкуба.

3. ПРОТОКОЛ САМОМАРШРУТИЗАЦИИ В ТРЕХМЕРНОМ ГИПЕРКУБЕ

Самомаршрутизация в трехмерном гиперкубе осуществляется как червячная маршрутизация путем прокладки бесконфликтного прямого пути между любыми источником и приемником через промежуточные коммутаторы. Она реализуется путем послышки пилотного пакета, содержащего маршрутную информацию, с фиксацией прямого пути в промежуточных коммутаторах. Пакет данных передается по прямому пути без задержек на его буферизацию в промежуточных узлах.

Сначала представим алгоритм в терминах дуг измерений X , Y и Z , считая, что исходный маршрут имеет вид $X_O Y_M Z_I$.

Алгоритм динамической самомаршрутизации (общая схема)

1. Если дуга X_O отсутствует ($X_O = 0$), то путь прокладывается по локальной дуге в узле s от источника к коммутатору. Переход к п. 3.

Если дуга X_O присутствует ($X_O \neq 0$), то путь прокладывается от источника в узле s по дуге X_O к коммутатору узла t . Переход к п. 2.

2. Если дуга Y_M отсутствует ($Y_M = 0$) и дуга Z_I отсутствует ($Z_I = 0$), то путь прокладывается по локальной дуге в узле t от коммутатора к источнику. Конец алгоритма.

3. Если дуга Y_M отсутствует ($Y_M = 0$), а дуга Z_I присутствует ($Z_I \neq 0$), то путь прокладывается по дуге Z_I от коммутатора узла t к приемнику в узле d . Конец алгоритма.

Если дуга Y_M присутствует ($Y_M \neq 0$), а дуга Z_I отсутствует ($Z_I = 0$), то путь прокладывается по дуге Y_I от коммутатора узла t к приемнику в узле d . Конец алгоритма.

Если дуга Y_M присутствует ($Y_M \neq 0$) и дуга Z_I присутствует ($Z_I \neq 0$), то проверяется наличие конфликта на дуге Y_M .

Если конфликта нет, то путь прокладывается по дуге Y_M от коммутатора узла t к коммутатору узла i . Переход к п. 4.

Если на дуге Y_M конфликт имеет место, то путь прокладывается по уникальной для узла дуге Y_D (разной для разных путей) от коммутатора узла t к коммутатору узла h . Переход к п. 5.

4. Путь прокладывается по дуге Z_I от коммутатора узла i к абоненту в узле-приемнике d . Конец алгоритма.

5. Проверяется наличие конфликта на дуге Y_C .

Если конфликта нет, то путь прокладывается по дуге Y_C от коммутатора узла h к коммутатору узла i . Переход к п. 4.

Если же конфликт имеет место, то переход к п. 6.

6. Путь прокладывается по дуге $Z_M = Z_I$ от коммутатора узла h к коммутатору узла e . Переход к п. 7.

7. Путь прокладывается по дуге $Y_I = Y_C$ от коммутатора узла e к приемнику в узле d . Конец алгоритма.

Теперь представим этот алгоритм в терминах путевой информации, которая содержится в пилотном пакете для червячной маршрутизации.

Путевую информацию будем представлять номерами выходных дуг в каждом узле, которые должны составлять прямой путь. Дуги нумеруются для каждого измерения в порядке возрастания их

длин числами от 0 до $p - 1$. Номер 0 задает либо внутриузловую дугу в узле-источнике и в узле-приемнике, либо отсутствие дуги между коммутаторами разных узлов. Таким образом, значения дуг m_1 измерения X задаются как числа m_1 ($0 \leq m_1 \leq p - 1$), значения дуг m_2 измерения Y задаются как числа m_2 ($0 \leq m_2 \leq p - 1$) и значения дуг m_3 измерения Z задаются как числа m_3 ($0 \leq m_3 \leq p - 1$).

Любой кратчайший путь задается тремя числами m_1 , m_2 и m_3 . Число m_1 задает дугу $X_O(m_1)$ от узла-источника s в промежуточный узел t . При $m_1 = 0$ узлы s и t совпадают. Число m_2 задает дугу $Y_M(m_2)$ от узла t к промежуточному узлу i . При $m_2 = 0$ узлы t и i совпадают. Число m_3 задает дугу $Z_I(m_3)$ от коммутатора узла i к узлу-приемнику d . Для прокладки прямого пути пилотный пакет должен содержать четыре числа m_1 , m_2 , m^* и m_3 , где $m^* = (m_1 - m_2) \bmod p$. Эти числа для каждого кратчайшего пути вычисляются каждым источником заранее.

Прямой путь прокладывается по следующему алгоритму.

*Алгоритм динамической самонашрутизации
(конкретная реализация)*

1. Если $m_1 = 0$, то путь прокладывается по локальной дуге в узле s от источника к коммутатору. Переход к п. 3.

Если $m_1 > 0$, то путь прокладывается от источника в узле s по дуге $X_O(m_1)$ к коммутатору узла t . Переход к п. 2.

2. Если $m_2 = 0$ и $m_3 = 0$, то путь прокладывается по локальной дуге в узле t от коммутатора к источнику. Конец алгоритма.

3. Если $m_2 = 0$ и $m_3 \neq 0$, то путь прокладывается по дуге $Z_I(m_3)$ от коммутатора узла t к приемнику в узле d . Конец алгоритма.

Если $m_2 \neq 0$ и $m_3 = 0$, то путь прокладывается по дуге $Y_I(m_2)$ от коммутатора узла t к приемнику в узле d . Конец алгоритма.

Если $m_2 > 0$, то проверяется наличие конфликта на дуге $Y_M(m_2)$. Если конфликта нет, то путь прокладывается по дуге $Y_M(m_2)$ от коммутатора узла t к коммутатору узла i . Переход к п. 4.

Если $m_2 > 0$ и имеет место конфликт, то путь прокладывается по дуге $Y_D(m_1)$ от коммутатора узла t к коммутатору узла h . Переход к п. 5.

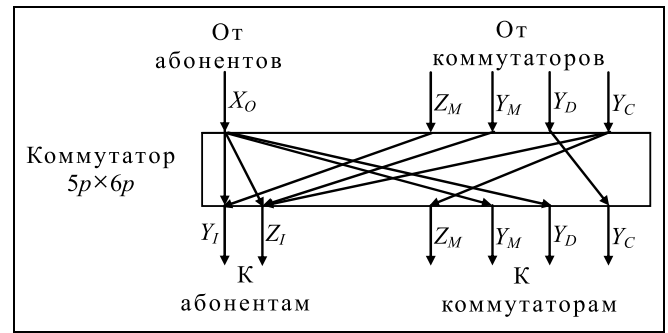


Рис. 8. Коммутатор узла неблокируемого трехмерного обобщенного гиперкуба

4. Путь прокладывается по дуге $Z_I(m_3)$ от коммутатора узла i к абоненту в узле-приемнике d . Конец алгоритма.

5. Проверяется конфликтность пути по дуге $Y_C(m^*)$.

Если конфликта нет, то путь прокладывается по дуге $Y_C(m^*)$ от коммутатора узла h к коммутатору узла i . Переход к п. 4.

Если же конфликт имеет место, то переход к п. 6.

6. Путь прокладывается по дуге $Z_M(m_3)$ от коммутатора узла h к коммутатору узла e . Переход к п. 7.

7. Путь прокладывается по дуге $Y_I(m^*)$ от коммутатора узла e к приемнику в узле d . Конец алгоритма.

Комментарий

- Прокладка пути в п. 3 из узла t по дуге $Y_D(m_1)$ гарантирует, что все дуги, проходимые из этого узла, оказываются различными, поскольку различны все входные дуги $X_O(m_1)$ в узел t .

- Прокладка пути в п. 5 из узла h по дуге $Y_C(m^*)$ гарантирует возврат пути в узел i , в который невозможно было попасть в п. 3 из-за конфликта.

Отметим, что прокладка прямых путей по приведенному алгоритму может содержать до четырех этапов. Однако передача всех пакетов данных по прямым путям осуществляется за один скачок. Последнее свойство позволяет утверждать, что неблокируемый гиперкуб практически (по задержкам передачи) имеет единичный диаметр.

Для реализации алгоритма самонашрутизации необходимо применять в каждом узле коммутатор $5p \times 6p$, схема которого представлена на рис. 8. Он не является полным коммутатором, так как используется только часть связей между входами и выходами. Поэтому его сложность можно оценить как $9p^2$ точек коммутации.

ЗАКЛЮЧЕНИЕ

Предложена новая структура трехмерного обобщенного гиперкуба на базе двумерных неблокируемых гиперкубов с минимальным числом дуг. Цель ее разработки состояла в обеспечении возможности прокладки бесконфликтных путей посредством их самомаршрутизации в узлах гиперкуба. Данная структура позволила исключить конфликты двухэтапных и трехэтапных путей и разработать алгоритм самомаршрутизации последних.

Предложенная структура и разработанный алгоритм обеспечивают неблокируемость гиперкуба, т. е. возможность передачи пакетов данных по прямым путям без задержки на их буферизацию в промежуточных узлах. Построенный гиперкуб имеет семерной набор $(p - 1)$ -й дуг, которые неравномерно распределены по его измерениям — одинарный набор в измерении X , четверной набор в измерении Y и двойной набор в измерении Z .

Рассмотренный двумерный обобщенный гиперкуб имеет топологию квазиполного орграфа. Такую же топологию имеет и двумерное полное мультикольцо. Они задают неблокируемые сети с минимальным числом дуг. Однако такие же сети могут иметь и топологию квазиполного графа, которая позволяет разминивать число узлов на число разных прямых путей между узлами. Последнее свойство обеспечивает встроенную отказоустойчивость сети.

Одна из конечных целей автора заключалась в применении сетей с топологией квазиполного графа в качестве граней трехмерного обобщенного гиперкуба для придания ему свойства отказоус-

тойчивости. Однако эта топология не имеет деления дуг и узлов по измерениям. Поэтому вопрос об ее применении остается открытым.

С другой стороны, многомерное полное мультикольцо имеет очень простое графическое представление — кольцо с разными диагоналями из каждого узла. Простота представления очень облегчает исследование его свойств и открывает возможность построения не только трехмерного неблокируемого мультикольца, но и дальнейшего повышения его размерности.

ЛИТЕРАТУРА

1. Bhuyan L.N. and Agrawal D.P. Generalized Hypercube and Hyperbus Structures for a Computer Network // IEEE Trans. on Computers. — 1984. — Vol. C-33, N 4. — P. 323—333.
2. Alverson R., Roweth D., Kaplan L., and Roweth D. Cray XC[®] Series Network. — URL: <https://www.cray.com/sites/default/files/resources/CrayXCNetwork.pdf> (дата обращения: 12.10.2017).
3. Подлазов В.С. p - p -перестраиваемость и отказоустойчивость двоекных p -ичных мультиколец и обобщенных гиперкубов // Автоматика и телемеханика. — 2002. — № 7. — С. 138—148.
4. Каравай М.Ф., Подлазов В.С. Распределенный полный коммутатор как «идеальная» системная сеть для многопроцессорных вычислительных систем // Управление большими системами. — 2011. — Вып. 34. — С. 92—116.
5. Scott S., Abts D., Kim J. and Dally W. The black widow high-radix Clos network // Proc. of 33rd Intern. Symp. Comp. Arch. (ISCA'2006). — URL: http://cva.stanford.edu/publications/2006/ISCA_YARC.pdf (дата обращения: 12.10.2017).

Статья представлена к публикации членом редсовета чл.-корр. РАН П.П. Пархоменко.

Подлазов Виктор Сергеевич — д-р техн. наук, гл. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ podlazov@ipu.ru.



Содержание сборника «Управление большими системами»

2018, вып. 72

Бибиков П.В., Стрельцова И.С. О контактном интегральном инварианте Калаби. — С. 27—32.

Васильев И.И., Чеботарев П.Ю. Об одном подходе к типизации ученых по библиометрическим данным. — С. 138—195.

Глуценко А.И. О методе определения скорости обучения нейронной сети для задачи оперативной настройки линейных регуляторов при управлении нелинейными объектами. — С. 52—107.

Горелов М.А. Принцип «Value at Risk» в иерархической игре. — С. 6—26.

Жожишвили А.В. Эквивалентность двух способов упорядочения множества морфизмов с общим источником. — С. 33—51.

Малышенко В.А., Малышенко К.А. Моделирование финансовой устойчивости предприятия в рамках фундаментального анализа развивающегося регионального фондового рынка. — С. 196—216.

Самарцев А.А., Иващенко В.А., Резников А.Ф. Мультиагентная модель процесса эвакуации людей из помещений при возникновении чрезвычайных ситуаций. — С. 217—244.

Федин Г.Г. Применение робастного подхода к задаче размещения транспортных узлов с подъездными путями в географическом регионе с существующей транспортной системой. — С. 108—137.

Тексты статей в свободном доступе на сайте <http://ubs.mtas.ru/>

МЕТОД ПОСТРОЕНИЯ КОНТРОЛИРУЮЩЕГО ТЕСТА ДЛЯ ЦИФРОВЫХ СИСТЕМ СО СТРУКТУРОЙ МИНИМАЛЬНОГО КВАЗИПОЛНОГО ГРАФА (НА ПРИМЕРЕ ГРАФА РАЗМЕРА 11×11)

В.А. Ведешенков

Разработан новый метод построения контролирующего теста для цифровых систем со структурой минимального квазиполного графа. Частные тесты сформированы для n фрагментов цифровой системы, каждый из которых состоит из коммутатора и $2m$ линий связи. Затем проверки частных тестов распределены в n групп с общими проверяющими абонентами. Множество из nm проверок представляет полный минимальный тест для одиночных неисправных компонентов цифровых систем рассматриваемой структуры. Приведен пример построения контролирующего теста для цифровой системы размера 11×11 . Отмечено, что предложенный метод можно применять для анализа цифровых систем рассматриваемой структуры с другими параметрами графа.

Ключевые слова: цифровая система, фрагмент, абонент, коммутатор, линия связи, минимальный квазиполный граф, контролирующий тест, частный тест.

ВВЕДЕНИЕ

В настоящей статье под цифровыми системами (ЦС) понимаются модели многомашинных или многопроцессорных вычислительных систем, отражающие необходимые диагностические свойства и параметры анализируемых вычислительных систем.

Минимальный квазиполный граф образуется на основе однородного двудольного графа, одну долю которого составляют коммутаторы $m \times m$, а другую — m -портовые абоненты. В одной доле имеется n коммутаторов, а в другой — n абонентов. Для рассматриваемых топологий параметры n и m связаны соотношением $n = m(m - 1)/\sigma + 1$ и не могут быть взяты произвольно [1]. Значение m выбирается минимальным, при котором любые два узла в одной доле связаны σ путями длины два через разные узлы в другой доле. Каждый такой путь проходит через один коммутатор, и разные пути проходят через разные коммутаторы.

В число возможных областей применения графов с подобной новой структурой входят отказоустойчивые многомашинные вычислительные системы реального времени, где, например, под-

множество вершин одной доли представляет совокупность процессорных элементов или вычислительных машин, а подмножество вершин другой доли — совокупность коммутаторов [1, 2]. Для систем ответственного применения время их восстановления часто играет решающую роль, а потому должно быть по возможности минимальным. Выход известен — применение систем с автоматическим восстановлением работоспособности. Трудности построения подобных систем существенно возрастают с увеличением числа компонентов, влияние неисправностей которых нужно устранить автоматически. В связи с этим представляют безусловный интерес системы с автоматическим восстановлением работоспособности после отказа одного из компонентов системы [3]. Поскольку физический ремонт неисправного компонента не может производиться, то можно говорить о самовосстановлении работоспособности ЦС.

Для построения таких самовосстанавливаемых ЦС представляется безусловно полезным фрагментный подход, в котором необходимые процедуры (в том числе диагностические процедуры) выполняются над частью (фрагментом) ЦС [4, 5]. Обязательной частью фрагментного подхода [4] является применение диагностического монитора —

Таблица 1

Таблица взаимодействия компонентов ЦС размера 11×11

A_j	A_k										
	1	2	3	4	5	6	7	8	9	10	11
1	∅	13 14	14 16	16 19	12 16	13 19	14 19	12 19	13 16	12 14	12 13
2	13 14	∅	14 15	15 17	17 20	13 17	14 20	15 20	13 20	14 17	13 15
3	14 16	14 15	∅	15 16	16 18	18 21	14 18	15 21	16 21	14 21	15 18
4	16 19	15 17	15 16	∅	16 17	17 19	19 22	15 19	16 22	17 22	15 22
5	12 16	17 20	16 18	16 17	∅	17 18	18 20	12 20	16 20	12 17	12 18
6	13 19	13 17	18 21	17 19	17 18	∅	18 19	19 21	13 21	17 21	13 18
7	14 19	14 20	14 18	19 22	18 20	18 19	∅	19 20	20 22	14 22	18 22
8	12 19	15 20	15 21	15 19	12 20	19 21	19 20	∅	20 21	12 21	12 15
9	13 16	13 20	16 21	16 22	16 20	13 21	20 22	20 21	∅	21 22	13 22
10	12 14	14 17	14 21	17 22	12 17	17 21	14 22	12 21	21 22	∅	12 22
11	12 13	13 15	15 18	15 22	12 18	13 18	18 22	12 15	13 22	12 22	∅

Требуется разработать способ построения контролирующего теста для ЦС со структурой минимального квазиполного графа (на примере графа с параметрами $n = 11$, $m = 5$, $\sigma = 2$).

2. СПОСОБ ПОСТРОЕНИЯ КОНТРОЛИРУЮЩЕГО ТЕСТА ДЛЯ ЦИФРОВЫХ СИСТЕМ СО СТРУКТУРОЙ МИНИМАЛЬНОГО КВАЗИПОЛНОГО ГРАФА (НА ПРИМЕРЕ ГРАФА С ПАРАМЕТРАМИ $N = 11$, $M = 5$, $\Sigma = 2$)

Реализация фрагментного способа [5] основана на следующих положениях:

— неисправный компонент K_x не влияет на работоспособность других компонентов ЦС, а потому фактическое техническое состояние компонента K_x (исправен или подозревается в неисправности) можно определить только по результату той проверки, в выполнении которой участвует компонент K_x ;

— предположение об одном неисправном компоненте вследствие редкого потока их отказов поз-

внешнего устройства для анализируемой ЦС, что затрудняет применение данного подхода в самовосстанавливаемых ЦС. В связи с этим более перспективным представляется фрагментный способ восстановления работоспособности ЦС без применения диагностического монитора [5].

Согласно этому способу с помощью контролирующего теста поочередно тестируется техническое состояние компонентов одного фрагмента ЦС. В зависимости от результатов тестирования: «работоспособен — не работоспособен» проверяющий абонент данного фрагмента выбирает направление последующих действий:

— продолжить контроль ЦС — при успешном результате тестирования;

— или перейти к устранению влияния неисправного компонента — при неудачном результате тестирования. Для этого проверяющий абонент идентифицирует компонент или группу компонентов, подозреваемых в неисправности. Затем выполняются подпрограммы реконфигурации, передающие исполнение функций неисправного компонента соответствующим исправным компонентам.

Одинаковая структура и состав фрагментов позволяют применять методы и средства восстановления, разработанные для одного фрагмента, для устранения влияния неисправного компонента в других фрагментах.

В данной статье предлагается регулярный метод построения контролирующего теста для ЦС со структурой минимального квазиполного графа; предполагается, что построенный контролирующий тест может быть использован в самовосстанавливаемых системах (например, рассмотренных в работе [5]).

1. ПОСТАНОВКА ЗАДАЧИ

Исходная ЦС имеет структуру минимального квазиполного графа $S_{n,m,\sigma}$, где n — число вершин (абонентов и коммутаторов) в каждой доле графа, m — число портов каждого коммутатора, $\sigma = 2$ — число путей между каждой парой абонентов, проходящих через разные коммутаторы. Структура конкретной ЦС представляется в табличном виде. (Таблица взаимодействия компонентов ЦС, рассматриваемой в качестве примера, показана в виде табл. 1, см. далее.)

Будем считать, что допускаются устойчивые редкие отказы одного из компонентов: абонента, коммутатора, линии связи, причем неисправности компонентов таковы, что прекращают работу тех компонентов, в которых они возникли, и не влияют на работоспособность других компонентов, не имеющих непосредственных связей с данным неисправным компонентом.



воляет считать, что исправны все компоненты ЦС, кроме одного из тех, которые входят в состав фрагмента, в котором обнаружен неисправный компонент. Это означает, что действия (операции), выполняемые этими исправными компонентами ЦС, являясь правильными, а их результаты достоверны;

— анализируемый фрагмент Φ_j включает в себя абонента $A_{k,2}$, коммутатор $C_{s,1}$ и линии связи $l(a_{j,1} - c_{s,1})$, $l(c_{s,1} - a_{k,2})$. Проверкой называется тестовая операция над компонентами фрагмента Φ_j , выполняемая следующим образом.

Исправный проверяющий абонент $A_{j,1}$ передает абоненту $A_{k,2}$ через коммутатор $C_{s,1}$ и соответствующие линии связи запрос (проверку) P_j . Абонент $A_{k,2}$, получив такое сообщение, подтверждает его получение абоненту $A_{j,1}$, который запоминает такой результат как оценку $r_j = 0$. Если в течение интервала Δt такого подтверждения от абонента $A_{k,2}$ не приходит, то абонент $A_{j,1}$ запоминает этот результат как оценку $r_j = 1$. Эти оценки используются абонентом $A_{j,1}$ для выбора последующих тестовых процедур.

В выполнении проверки P_j участвуют компоненты: проверяющий абонент $A_{j,1}$, линия связи $l(a_{j,1} - c_{s,1})$, коммутатор $C_{s,1}$, линия связи $l(c_{s,1} - a_{k,2})$, проверяемый абонент $A_{k,2}$.

Без названий компонентов состав проверки выглядит так:

$$P_j = A_{j,1}, l(a_{j,1} - c_{s,1}), C_{s,1}, l(c_{s,1} - a_{k,2}), A_{k,2}.$$

Как видно из этой записи, индексы линий связи однозначно определяются индексами компонентов, которые они соединяют. Для уменьшения длины записи исключим из нее символы линий связи, получим обозначение проверки, в котором показаны только символы «активных» компонентов (абонент, коммутатор):

$$P_j = A_{j,1} \rightarrow C_{s,1} \rightarrow A_{k,2}.$$

В рассматриваемом далее примере ЦС каждый «активный» компонент имеет свой номер, поэтому при записи проверок можно обойтись без символов компонентов, а писать только их номера. Например, далее в табл. 4 найдем такие проверки:

$$P_1 = 1 \rightarrow 12 \rightarrow 5, \quad P_6 = 2 \rightarrow 13 \rightarrow 1, \\ P_{31} = 7 \rightarrow 14 \rightarrow 10 \text{ и др.}$$

Каждый абонент ЦС связан с каждым другим абонентом двумя линиями связи, проходящими через два разных коммутатора. Пути проверок, проходящие между абонентами $A_{j,1}$ и $A_{k,2}$ в одном и

том же направлении, но через разные коммутаторы $C_{s,1}$ и $C_{t,2}$, назовем сопряженными. Коммутаторы $C_{s,1}$ и $C_{t,2}$ также будем называть сопряженными относительно абонентов $A_{j,1}$ и $A_{k,2}$.

Структура взаимодействия компонентов анализируемой ЦС показана в табл. 1. Строки и столбцы табл. 1 обозначены номерами абонентов $A_j, A_k, j, k = 1, \dots, 11$, соответственно. В клетке на пересечении j -й строки и k -го столбца записаны номера коммутаторов $C_{s,1}$ и $C_{t,2}$, $s, t = 12, \dots, 22$, через которые проходят сопряженные пути для данной пары абонентов. Так как связи между абонентами двусторонние, то номера коммутаторов $C_{s,1}$ и $C_{t,2}$, связывающих абонента j -й строки с абонентом k -го столбца, совпадают с номерами коммутаторов $C_{s,1}$ и $C_{t,2}$, связывающих абонента k -й строки с абонентом j -го столбца. Например, в клетках $j = 3, k = 7$ и $j = 7, k = 3$ записаны номера коммутаторов 14 и 18, являющихся сопряженными для абонентов 3 и 7. Поскольку каждую пару абонентов связывает «индивидуальная» пара путей, то такая пара коммутаторов записана только в этих двух клетках. Знаки $\emptyset$ в диагональных клетках отражают тот факт, что абонент не взаимодействует сам с собой.

Основные требования к контролирующему тесту:

— каждая линия связи должна входить хотя бы в одну проверку, т. е. тестироваться не менее одного раза;

— в состав контролирующего теста должен входить только один из каждой пары сопряженных путей проверок, чтобы второй путь проверки можно было оставить в резерве.

В соответствии с записями табл. 1 в состав рассматриваемой ЦС входят одиннадцать абонентов с номерами 1, ..., 11, одиннадцать коммутаторов с номерами 12, ..., 22 и 110 линий связи между указанными абонентами и коммутаторами. Эти 110 линий связи объединены в 55 пар сопряженных путей проверок. Согласно упомянутым требованиям, в контролирующий тест должен входить только один путь из каждой пары сопряженных путей. Вариант контролирующего теста (см. работу [5]) был получен после нескольких попыток его построения потому, что выбор одного пути из каждой пары сопряженных путей зависит от ранее выбранных путей и приводит к перебору возможных вариантов.

Предлагаемая процедура построения контролирующего теста состоит из двух этапов:

— построение контролирующего теста для фрагмента F_i , состоящего из коммутатора C_i и десяти присоединенных к нему линий связи; назовем это подмножество проверок частным тестом коммута-

тора C_i (отметим, что по составу фрагмент F_i отличается от фрагмента Φ_j , упоминавшегося ранее);

— распределение проверок из n частных тестов для коммутаторов C_i в n групп с общим проверяющим абонентом $A_{j,1}$.

Такая двухэтапная процедура построения контролирующего теста приводит к существенному сокращению числа вариантов, перебираемых на втором этапе, благодаря их целенаправленному выбору.

2.1. Метод построения контролирующего теста для фрагментов цифровой системы со структурой минимального квазиполного графа (на примере графа с параметрами $n = 11$, $m = 5$, $\sigma = 2$)

Фрагмент F_i , состоящий из коммутатора C_i и десяти присоединенных линий связи, можно представить в виде звездообразного графа: центральная вершина (коммутатор C_i) и 5 пар дуг (линий связи), соединяющих центр с периферийными вершинами (абонентами $A_{j,1}$, ..., $A_{j,5}$). Каждая проверка компонентов фрагмента F_i отображается на графе путем из одной вершины в другую, проходящим через центр и две дуги, соединяющие центр с этими вершинами. Для построения частного теста фрагмента F_i нужно обойти все вершины графа,

например, в порядке увеличения номеров вершин. В процессе такого обхода в одной проверке P_j будет протестирована линия связи от абонента A_j к коммутатору C_i , а в другой проверке P_k — инверсная линия связи от коммутатора C_i к абоненту A_j . После обхода всех вершин графа по одному разу будут протестированы все линии связи, прилегающие к коммутатору C_i . Таким образом, частный тест является полным для неисправных линий связи. Частный тест является минимальным, так как каждая линия связи входит только в одну проверку и потому тестируется только один раз.

Построив частные тесты для всех коммутаторов ЦС, получим множество проверок, необходимое и достаточное для построения полного контролирующего теста ЦС.

В табл. 2 показана реализация двух вариантов частных тестов для коммутаторов C_i , $i = 12, \dots, 22$. При ее построении приняты такие обозначения: левый столбец показывает номера коммутаторов C_i , нижний индекс которого (1 или 2) отражает номер рассматриваемого варианта частных тестов (о них подробнее будет сказано далее); второй столбец содержит номера абонентов, связанных с коммутатором C_i , причем порядок записи номеров отражает порядок их последующего обхода при пост-

Таблица 2

Таблица состава частных тестов для цифровой системы размера 11×11

Коммутатор C_i	Абоненты $A_{j,1}, A_{j,2}, A_{j,3}, A_{j,4}, A_{j,5}$	Пути проверок для коммутатора C_i
12_1	1, 5, 8, 10, 11	$1 \rightarrow 12 \rightarrow 5; 5 \rightarrow 12 \rightarrow 8; 8 \rightarrow 12 \rightarrow 10; 10 \rightarrow 12 \rightarrow 11; 11 \rightarrow 12 \rightarrow 1$
13_1 13_2	1, 2, 6, 9, 11 11, 9, 6, 2, 1	$1 \rightarrow 13 \rightarrow 2; 2 \rightarrow 13 \rightarrow 6; 6 \rightarrow 13 \rightarrow 9; 9 \rightarrow 13 \rightarrow 11; 11 \rightarrow 13 \rightarrow 1$ $11 \rightarrow 13 \rightarrow 9; 9 \rightarrow 13 \rightarrow 6; 6 \rightarrow 13 \rightarrow 2; 2 \rightarrow 13 \rightarrow 1; 1 \rightarrow 13 \rightarrow 11$
14_1	1, 2, 3, 7, 10	$1 \rightarrow 14 \rightarrow 2; 2 \rightarrow 14 \rightarrow 3; 3 \rightarrow 14 \rightarrow 7; 7 \rightarrow 14 \rightarrow 10; 10 \rightarrow 14 \rightarrow 1$
15_1 15_2	2, 3, 4, 8, 11 11, 8, 4, 3, 2	$2 \rightarrow 15 \rightarrow 3; 3 \rightarrow 15 \rightarrow 4; 4 \rightarrow 15 \rightarrow 8; 8 \rightarrow 15 \rightarrow 11; 11 \rightarrow 15 \rightarrow 2$ $11 \rightarrow 15 \rightarrow 8; 8 \rightarrow 15 \rightarrow 4; 4 \rightarrow 15 \rightarrow 3; 3 \rightarrow 15 \rightarrow 2; 2 \rightarrow 15 \rightarrow 11$
16_1	1, 3, 4, 5, 9	$1 \rightarrow 16 \rightarrow 3; 3 \rightarrow 16 \rightarrow 4; 4 \rightarrow 16 \rightarrow 5; 5 \rightarrow 16 \rightarrow 9; 9 \rightarrow 16 \rightarrow 1$
17_1 17_2	2, 4, 5, 6, 10 10, 6, 5, 4, 2	$2 \rightarrow 17 \rightarrow 4; 4 \rightarrow 17 \rightarrow 5; 5 \rightarrow 17 \rightarrow 6; 6 \rightarrow 17 \rightarrow 10; 10 \rightarrow 17 \rightarrow 2$ $10 \rightarrow 17 \rightarrow 6; 6 \rightarrow 17 \rightarrow 5; 5 \rightarrow 17 \rightarrow 4; 4 \rightarrow 17 \rightarrow 2; 2 \rightarrow 17 \rightarrow 10$
18_1	3, 5, 6, 7, 11	$3 \rightarrow 18 \rightarrow 5; 5 \rightarrow 18 \rightarrow 6; 6 \rightarrow 18 \rightarrow 7; 7 \rightarrow 18 \rightarrow 11; 11 \rightarrow 18 \rightarrow 3$
19_1 19_2	1, 4, 6, 7, 8 8, 7, 6, 4, 1	$1 \rightarrow 19 \rightarrow 4; 4 \rightarrow 19 \rightarrow 6; 6 \rightarrow 19 \rightarrow 7; 7 \rightarrow 19 \rightarrow 8; 8 \rightarrow 19 \rightarrow 1$ $8 \rightarrow 19 \rightarrow 7; 7 \rightarrow 19 \rightarrow 6; 6 \rightarrow 19 \rightarrow 4; 4 \rightarrow 19 \rightarrow 1; 1 \rightarrow 19 \rightarrow 8$
20_1	2, 5, 7, 8, 9	$2 \rightarrow 20 \rightarrow 5; 5 \rightarrow 20 \rightarrow 7; 7 \rightarrow 20 \rightarrow 8; 8 \rightarrow 20 \rightarrow 9; 9 \rightarrow 20 \rightarrow 2$
21_1 21_2	3, 6, 8, 9, 10 10, 9, 8, 6, 3	$3 \rightarrow 21 \rightarrow 6; 6 \rightarrow 21 \rightarrow 8; 8 \rightarrow 21 \rightarrow 9; 9 \rightarrow 21 \rightarrow 10; 10 \rightarrow 21 \rightarrow 3$ $10 \rightarrow 21 \rightarrow 9; 9 \rightarrow 21 \rightarrow 8; 8 \rightarrow 21 \rightarrow 6; 6 \rightarrow 21 \rightarrow 3; 3 \rightarrow 21 \rightarrow 10$
22_1 22_2	4, 7, 9, 10, 11 4, 7, 9, 11, 10	$4 \rightarrow 22 \rightarrow 7; 7 \rightarrow 22 \rightarrow 9; 9 \rightarrow 22 \rightarrow 10; 10 \rightarrow 22 \rightarrow 11; 11 \rightarrow 22 \rightarrow 4$ $4 \rightarrow 22 \rightarrow 7; 7 \rightarrow 22 \rightarrow 9; 9 \rightarrow 22 \rightarrow 11; 11 \rightarrow 22 \rightarrow 10; 10 \rightarrow 22 \rightarrow 4$

Таблица 3

Таблица сопряженных путей проверок

№ п/п	Проверки
1	$P_{1,1} = 1 \rightarrow 13 \rightarrow 2$ $P_{1,2} = 1 \rightarrow 14 \rightarrow 2$
2	$P_{2,1} = 2 \rightarrow 14 \rightarrow 3$ $P_{2,2} = 2 \rightarrow 15 \rightarrow 3$
3	$P_{3,1} = 3 \rightarrow 15 \rightarrow 4$ $P_{3,2} = 3 \rightarrow 16 \rightarrow 4$
4	$P_{4,1} = 4 \rightarrow 16 \rightarrow 5$ $P_{4,2} = 4 \rightarrow 17 \rightarrow 5$
5	$P_{5,1} = 5 \rightarrow 17 \rightarrow 6$ $P_{5,2} = 5 \rightarrow 18 \rightarrow 6$
6	$P_{6,1} = 6 \rightarrow 18 \rightarrow 7$ $P_{6,2} = 6 \rightarrow 19 \rightarrow 7$
7	$P_{7,1} = 7 \rightarrow 19 \rightarrow 8$ $P_{7,2} = 7 \rightarrow 20 \rightarrow 8$
8	$P_{8,1} = 8 \rightarrow 20 \rightarrow 9$ $P_{8,2} = 8 \rightarrow 21 \rightarrow 9$
9	$P_{9,1} = 9 \rightarrow 21 \rightarrow 10$ $P_{9,2} = 9 \rightarrow 22 \rightarrow 10$
10	$P_{10,1} = 10 \rightarrow 22 \rightarrow 11$ $P_{10,2} = 10 \rightarrow 12 \rightarrow 11$
11	$P_{11,1} = 11 \rightarrow 12 \rightarrow 1$ $P_{11,2} = 11 \rightarrow 13 \rightarrow 1$

роении частного теста; в третьем столбце показаны пути пяти проверок, образующих частный тест для коммутатора C_p , $i = 12, \dots, 22$.

Пути проверок, записанные в третьем столбце, образованы достаточно просто: между двумя смежными номерами абонентов из второго столбца вставляется номер коммутатора из первого столбца. Двигаясь от номера левого абонента вправо по клетке второго столбца, построим поочередно пять проверок, образующих частный тест для коммутатора с номером из клетки первого столбца.

Первый вариант контролирующего теста ЦС составлен из частных тестов, пути проверок которых построены в порядке увеличения номеров абонентов. Сравнение записей этого варианта выявило одиннадцать пар сопряженных путей проверок, показанных в табл. 3. Появление таких пар сопряженных путей проверок не случайно, оно является следствием структурных особенностей анализируемой ЦС, а именно, наличия двух путей между любой парой абонентов. Действительно, в табл. 1 нетрудно найти клетки, в которых находятся пары с четным и нечетным номерами коммутаторов, часть этих пар показана в табл. 3. Например, коммутаторы 13, 14 в клетках 1, 2 и 2, 1 табл. 1 входят в состав сопряженных проверок $P_{1,1}$ и $P_{1,2}$ в табл. 3, коммутаторы 14, 15 в клетках 2, 3 и 3, 2 табл. 1 — в состав проверок $P_{2,1}$ и $P_{2,2}$ в табл. 3 и т. д., ..., коммутаторы 21, 22 в проверках $P_{9,1}$ и $P_{9,2}$, коммутаторы 12, 13 в проверках $P_{11,1}$ и $P_{11,2}$.

Для того чтобы в состав контролирующего теста ЦС не входили пары сопряженных путей проверок, правило построения проверок в порядке увеличения номеров абонентов заменим на два таких правила:

- для коммутаторов с четными номерами пути проверок частных тестов строятся в порядке увеличения номеров абонентов,

- для коммутаторов с нечетными номерами пути проверок частных тестов строятся в обратном порядке: в порядке уменьшения номеров абонентов.

Второе правило применено для построения частных тестов для коммутаторов с номерами 13, 15, 17, 19, 21. Полученные таким образом частные тесты показаны в строках табл. 2, отмеченных номерами коммутаторов с индексом 2. Нетрудно проверить, что инвертирование направления путей проверок в частных тестах для коммутаторов с нечетными номерами исключило десять из одиннадцати пар сопряженных путей проверок, показанных в табл. 3. Второе правило построения частных тестов не повлияло на частные тесты для коммутаторов 22 и 12 (естественно, это же четные номера). Поэтому в составе показанных частных тестов осталась десятая пара сопряженных комму-

таторов 22, 12 в проверках $P_{10,1} = 10 \rightarrow 22 \rightarrow 11$ и $P_{10,2} = 10 \rightarrow 12 \rightarrow 11$. Дело в том, что номер 22 — четный, а благодаря нечетному числу абонентов ($n = 11$) номер 22 в частных тестах соседствует с другим четным номером 12. Для исключения этой пары сопряженных проверок сделаем такую коррекцию: изменим на инверсное направление путь проверки ($10 \rightarrow 22 \rightarrow 11$) и скорректируем те пути проверок для коммутатора 22, которые выполняются до и после этой проверки. В результате таких действий получим частный тест, показанный в строке 22₂.

Таким образом, два предложенных правила построения частных тестов хорошо работают для ЦС с четным числом коммутаторов n , но требуют дополнительной коррекции одного частного теста при нечетном числе коммутаторов.

2.2. Метод построения контролирующего теста для цифровых систем со структурой минимального квазиполного графа (на примере графа с параметрами $n = 11$, $m = 5$, $\sigma = 2$)

Множество путей проверок частных тестов для коммутаторов с номерами 12₁, 14₁, 16₁, 18₁, 20₁ и 13₂, 15₂, 17₂, 19₂, 21₂, 22₂, показанных в табл. 2, яв-

ляется необходимым и достаточным для построения контролирующего теста ЦС, удовлетворяющего заданным требованиям. Но для получения достоверных результатов их выполнения нужны исправные проверяющие абоненты $A_j, j = 1, \dots, 11$. Чтобы удовлетворить этому требованию, будем назначать очередную проверку P_j к выполнению только в том случае, если исправность проверяющего абонента $A_{j,1}$ подтверждена в ранее выполненных проверках.

Таблица 4

Таблица контролирующего теста для ЦС размера 11×11

B_j	Проверки $P_{j,1} = A_j \rightarrow C_i \rightarrow A_k$	B_j	Проверки $P_{j,1} = A_j \rightarrow C_i \rightarrow A_k$
1	$P_1 = 1 \rightarrow 12 \rightarrow 5$	31	$P_{31} = 7 \rightarrow 14 \rightarrow 10$
2	$P_2 = 1 \rightarrow 13 \rightarrow 11$	32	$P_{32} = 7 \rightarrow 18 \rightarrow 11$
3	$P_3 = 1 \rightarrow 14 \rightarrow 2$	33	$P_{33} = 7 \rightarrow 19 \rightarrow 6$
4	$P_4 = 1 \rightarrow 16 \rightarrow 3$	34	$P_{34} = 7 \rightarrow 20 \rightarrow 8$
5	$P_5 = 1 \rightarrow 19 \rightarrow 8$	35	$P_{35} = 7 \rightarrow 22 \rightarrow 9$
6	$P_6 = 2 \rightarrow 13 \rightarrow 1$	36	$P_{36} = 8 \rightarrow 12 \rightarrow 10$
7	$P_7 = 2 \rightarrow 14 \rightarrow 3$	37	$P_{37} = 8 \rightarrow 15 \rightarrow 4$
8	$P_8 = 2 \rightarrow 15 \rightarrow 11$	38	$P_{38} = 8 \rightarrow 19 \rightarrow 7$
9	$P_9 = 2 \rightarrow 17 \rightarrow 10$	39	$P_{39} = 8 \rightarrow 20 \rightarrow 9$
10	$P_{10} = 2 \rightarrow 20 \rightarrow 5$	40	$P_{40} = 8 \rightarrow 21 \rightarrow 6$
11	$P_{11} = 3 \rightarrow 14 \rightarrow 7$	41	$P_{41} = 9 \rightarrow 13 \rightarrow 6$
12	$P_{12} = 3 \rightarrow 15 \rightarrow 2$	42	$P_{42} = 9 \rightarrow 16 \rightarrow 1$
13	$P_{13} = 3 \rightarrow 16 \rightarrow 4$	43	$P_{43} = 9 \rightarrow 20 \rightarrow 2$
14	$P_{14} = 3 \rightarrow 18 \rightarrow 5$	44	$P_{44} = 9 \rightarrow 21 \rightarrow 8$
15	$P_{15} = 3 \rightarrow 21 \rightarrow 10$	45	$P_{45} = 9 \rightarrow 22 \rightarrow 11$
16	$P_{16} = 4 \rightarrow 15 \rightarrow 3$	46	$P_{46} = 10 \rightarrow 12 \rightarrow 11$
17	$P_{17} = 4 \rightarrow 16 \rightarrow 5$	47	$P_{47} = 10 \rightarrow 14 \rightarrow 1$
18	$P_{18} = 4 \rightarrow 17 \rightarrow 2$	48	$P_{48} = 10 \rightarrow 17 \rightarrow 6$
19	$P_{19} = 4 \rightarrow 19 \rightarrow 1$	49	$P_{49} = 10 \rightarrow 21 \rightarrow 9$
20	$P_{20} = 4 \rightarrow 22 \rightarrow 7$	50	$P_{50} = 10 \rightarrow 22 \rightarrow 4$
21	$P_{21} = 5 \rightarrow 12 \rightarrow 8$	51	$P_{51} = 11 \rightarrow 12 \rightarrow 1$
22	$P_{22} = 5 \rightarrow 16 \rightarrow 9$	52	$P_{52} = 11 \rightarrow 13 \rightarrow 9$
23	$P_{23} = 5 \rightarrow 17 \rightarrow 4$	53	$P_{53} = 11 \rightarrow 15 \rightarrow 8$
24	$P_{24} = 5 \rightarrow 18 \rightarrow 6$	54	$P_{54} = 11 \rightarrow 18 \rightarrow 3$
25	$P_{25} = 5 \rightarrow 20 \rightarrow 7$	55	$P_{55} = 11 \rightarrow 22 \rightarrow 10$
26	$P_{26} = 6 \rightarrow 13 \rightarrow 2$		
27	$P_{27} = 6 \rightarrow 17 \rightarrow 5$		
28	$P_{28} = 6 \rightarrow 18 \rightarrow 7$		
29	$P_{29} = 6 \rightarrow 19 \rightarrow 4$		
30	$P_{30} = 6 \rightarrow 21 \rightarrow 3$		

В табл. 4 показан вариант контролирующего теста для данной ЦС, построенный на основе приведенных проверок P_j из табл. 2. При формировании табл. 4 все 55 проверок частных тестов распределены в 11 групп, в каждой из которых 5 проверок имеют общий проверяющий абонент A_j (имеют один и тот же номер). Правила заполнения строк каждой группы просты: из табл. 3 выбираются те проверки, для которых абонент A_j является общим. Количество проверок в каждой группе равно значению m — числу портов у абонента и коммутатора. Упорядочивать проверки в каждой группе можно по разным критериям, например, в табл. 4 проверки записаны в порядке увеличения номеров коммутаторов.

Сформированная таким образом таблица контролирующего теста (см. табл. 4) содержит $R = n \times m = 55$ проверок P_j , что естественно совпадает с числом проверок в табл. 2, так как оно определяется числом линий связи в рассматриваемой ЦС.

Отметим, что порядок использования абонентов в качестве проверяющих, показанный в табл. 4, не является единственно возможным. Результат исполнения каждой проверки зависит от технического состояния ее проверяемых компонентов и проверяющего абонента A_j . Поэтому группы проверок с общим проверяющим абонентом A_j можно менять местами при условии, что исправность абонента A_j данной группы подтверждена ранее выполненными проверками. Так, в табл. 4 после исполнения первых 25 проверок будет протестирован хотя бы по одному разу каждый из 11 абонентов, в том числе абонент 6 проверкой P_{24} . Поэтому перестановка местами групп проверок с абонентами 6—11 никак не повлияет на достоверность результатов исполнения проверок с номерами 26—55 (см. табл. 4).

Напомним порядок выполнения контролирующего теста ЦС. Сначала выполняются проверки $P_1, \dots, P_5$, в которых проверяющим служит абонент 1. После успешного выполнения этих проверок абонент 1 передает управление абоненту 2 для выполнения проверок $P_6, \dots, P_{10}$. После успешного выполнения этих проверок абонент 2 передает управление абоненту 3 для тестирования проверок $P_{11}, \dots, P_{15}$ и т. д. Из записей табл. 4 видно, что исправность абонента 2 будет подтверждена проверкой P_3 , а исправность абонента 3 — проверкой P_4 и т. д. При неудачном результате выполнения любой проверки абонент A_j , зафиксировавший такой исход, переходит к диагностированию неисправного компонента и устранению его влияния на работоспособность ЦС [4, 5].



Проверка P_1 , для которой проверяющим служит абонент 1, может нарушить этот порядок: в случае неисправности абонента 1 система контроля остановится. Чтобы устранить влияние неисправности абонента 1, нужно запустить контролирующий тест с другого начального абонента, например, с абонента 2. Для такой оперативной замены начального абонента можно воспользоваться следующим вариантом действий.

При проектировании системы контроля ЦС организуем ее запуск по сигналам системных часов. Пусть запуск абонента 1 в качестве начального абонента контролирующего теста ЦС запланирован на момент времени T_0 , а запуск абонента 2 в качестве начального абонента контролирующего теста ЦС запланирован на момент времени $T_0 + \Delta T$, где интервал ΔT определен, например, экспериментально.

При успешном выполнении проверки P_1 абонент 1 передает абоненту 2 сигнал блокировки запуска системы контроля от системных часов (в момент времени $T_0 + \Delta T$) и таким образом предотвращает двойной запуск системы контроля ЦС. Если же проверка P_1 не выполнена, в том числе, и из-за неисправности абонента 1, то абонент 1 не отправит абоненту 2 команду на блокировку запуска системы контроля. В этом случае абонент 2 в момент времени $T_0 + \Delta T$ запустит систему контроля с проверки P_6 . В предположении, что в ЦС есть только один неисправный компонент, такая замена начального абонента позволит гарантированно запустить контролирующий тест, а в случае обнаружения неисправного компонента — перейти к этапам диагностирования и устранения влияния локализованного неисправного компонента, в том числе, и абонента 1.

Приведенные правила построения частных тестов для отдельных коммутаторов и контролирующего теста для всей ЦС сравнительно просто распространяются на ЦС с другими значениями параметров n и m . Действительно, частные тесты для одного коммутатора имеют m проверок, а контролирующий тест для ЦС распределяет проверки частных тестов n коммутаторов в n групп с общим проверяющим абонентом. Исключение составляют ЦС с нечетным значением n , для которых нужна коррекция частных тестов для одного коммутатора. Таким образом, контролирующий тест ЦС с параметрами n и m содержит $R = (n \times m)$ проверок.

ЗАКЛЮЧЕНИЕ

Разработан новый метод построения контролирующего теста для цифровых систем со структурой

минимального квазиполного графа. Частные тесты сформированы для n фрагментов ЦС, каждый из которых состоит из коммутатора и $2m$ линий связи. Затем проверки частных тестов распределены в n групп с общими проверяющими абонентами. Множество из nm проверок представляет полный минимальный тест для одиночных неисправных компонентов ЦС рассматриваемой структуры. Приведен пример построения контролирующего теста для ЦС размера 11×11 .

Предложенный метод построения контролирующего теста достаточно общий и может быть применен для анализа ЦС со структурой минимального квазиполного графа с параметрами графа, отличными от рассмотренного примера. Но параметры n и m должны удовлетворять соотношению $n = m(m - 1)/\sigma + 1$ [1]. Что касается параметра σ — числа путей между двумя абонентами, проходящих через σ коммутаторов, то этот параметр не влияет на процедуры построения частных тестов отдельных фрагментов. Но для диагностирования и устранения влияния неисправного компонента в ЦС необходимо значение $\sigma = 2$ [4, 5].

ЛИТЕРАТУРА

1. Каравай М.Ф., Подлазов В.С. Распределенный полный коммутатор как «идеальная» системная сеть для многопроцессорных вычислительных систем // Управление большими системами. — 2011. — Вып. 34. — С. 92—116. — URL: <http://ubs.mtas.ru/upload/library/UBS3405.pdf> (дата обращения 22.11.2017).
2. Alverson B., Froese E., Kaplan L., Roweth D. Cray XCTM Series Network, WP-Aries01-1112. — 2012. Cray Inc., 2012. — 28 p. — URL: <http://www.cray.com/sites/default/files/resources/CrayXCNetwork.pdf> (дата обращения: 24.11.2017).
3. Ведешенков В.А. Организация диагностирования одиночных неисправных компонентов цифровых систем со структурой минимального квазиполного графа размера 7×7 с двумя путями между двумя абонентами // Проблемы управления. — 2016. — № 1. — С. 65—72.
4. Ведешенков В.А. Подход к фрагментному диагностированию компонентов цифровых систем со структурой минимального квазиполного графа (на примере графа размера 7×7) // Проблемы управления. — 2016. — № 6. — С. 53—58.
5. Ведешенков В.А. Фрагментный способ восстановления работоспособности цифровых систем со структурой минимального квазиполного графа с двумя путями между двумя абонентами (на примере графа размера 11×11) // Автоматика и телемеханика. — 2018. — № 3. — С. 76—91.

Статья представлена к публикации членом редсовета чл.-корр. РАН П.П. Пархоменко.

Ведешенков Виктор Алексеевич — д-р техн. наук, гл. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва.

СОЗДАНИЕ ИНФОРМАЦИОННО-ТЕХНОЛОГИЧЕСКОГО РЕЗЕРВА В РАСПРЕДЕЛЕННЫХ СИСТЕМАХ ОБРАБОТКИ ДАННЫХ

С.К. Сомов

Представлены основные этапы формирования информационно-технологического резерва данных в распределенных автоматизированных информационно-управляющих системах. Формирование резерва выполняется на основе результатов анализа предметных областей пользователей, регулярных запросов к системе и анализа последовательности выполнения процедур обработки запросов. Поставлены задачи синтеза структуры резерва и оптимального размещения его копий по узлам компьютерной сети. Отмечено, что данный вид резервирования позволяет сократить время обработки типовых, регулярных запросов к системе благодаря тому, что резерв содержит заранее подготовленные массивы данных, используемые затем при обработке запросов.

Ключевые слова: распределенная система обработки данных, компьютерная сеть, информационно-технологический резерв.

ВВЕДЕНИЕ

Распределенные системы обработки данных (РСОД) различных типов, назначения и масштаба предназначены для обеспечения доступа огромному числу пользователей к программным, информационным и вычислительным ресурсам компьютерной сети. Это сложные технические системы, работа которых основана на взаимодействии большого количества аппаратных, программных и информационных компонентов. Распределенные системы объединяют компьютеры пользователей и информационные и программные ресурсы серверов компьютерной сети, удаленные друг от друга на сотни и тысячи километров.

Проектирование и создание таких систем — большой, ресурсоемкий и сложный процесс, в ходе которого необходимо решить множество важных задач. Одна из таких задач состоит в обеспечении высокой производительности системы. Для этого необходимо:

- применение качественных аппаратных средств (надежные каналы связи с большой пропускной способностью, высокопроизводительные серверы, надежные и быстродействующие устройства хранения информации);

- проектирование оптимальной по тем или иным критериям логической и физической струк-

туры распределенных по узлам сети данных, с которыми работают такие системы;

- размещение программных и информационных ресурсов максимально близко к пользователям, которым необходимы эти ресурсы.

Для приближения программных и информационных ресурсов к пользователям могут применяться методы [1—5], предусматривающие репликацию массивов данных и таблиц распределенных баз данных и распределение по нескольким узлам сети горизонтальных и/или вертикальных фрагментов таблиц распределенных баз данных.

Перечисленные решения позволяют увеличить производительность системы и надежность ее работы. Производительность системы увеличивается благодаря размещению данных в нескольких ближайших к пользователям узлах сети. Надежность работы системы также повышается, так как при сбое обработки запроса пользователя к данным, расположенным в одном узле сети, запрос можно переадресовать в другой работоспособный узел с необходимыми данными.

В настоящей статье предлагается новый эффективный метод повышения производительности и надежности работы распределенных систем, суть которого заключается в создании и оптимальном размещении в узлах компьютерной сети информационно-технологического резерва (ИТР).



1. ЦЕЛИ И ЭТАПЫ ФОРМИРОВАНИЯ ИНФОРМАЦИОННО-ТЕХНОЛОГИЧЕСКОГО РЕЗЕРВА

Информационно-технологическое резервирование развивает и обобщает методы структурно-технологической избыточности [6]. Информационно-технологический резерв основывается на использовании преимуществ:

- концепции репликации по узлам сети массивов данных и таблиц баз данных или их фрагментов [7—12], что обеспечивает повышение надежности и производительности распределенной системы;

- концепции предметно-ориентированных баз данных [13—15], структура данных которых построена с учетом результатов анализа предметных областей пользователей и их запросов к данным и оптимизирована в целях сокращения времени получения отчетов и результатов бизнес-анализа хранящихся в них данных.

1.1. Специфика информационно-технологического резервирования

Специфика метода информационно-технологического резервирования заключается в его нескольких особенностях. Применение ИТР основывается на том, что во многих случаях характеристики и особенности обработки регламентных, регулярных запросов пользователей к РСОД позволяют разбить их на несколько групп запросов различного типа. Для каждой такой группы можно выделить инвариантную этой группе общую часть технологии обработки запросов и подмножество используемых ими данных. Это означает, что в технологии обработки запросов группы есть общая часть, которая используется при обработке всех запросов данной группы. Эта часть технологии выделяется в виде подмножества процедур обработки данных, которая выполняется заранее, до поступления регламентных запросов. В результате работы этого подмножества процедур формируются промежуточные, вспомогательные массивы данных, которые затем используются при обработке запросов соответствующего типа. Очевидно, что изложенное выше за редким исключением не относится к узкоспециализированным, разовым ad-hoc запросам, а применимо к регулярным, регламентным запросам.

Таким образом, ИТР представляет собой множество процедур обработки исходных данных РСОД и сформированных ими массивов промежуточных данных, которые распределяются в виде копий по узлам компьютерной сети и затем используются при обработке регулярных запросов пользователей системы. Применение информационно-технологического резервирования целесооб-

разно и наиболее эффективно в автоматизированных информационно-управляющих системах, функционирующих на основе РСОД. Это объясняется, прежде всего, тем, что, значительная часть генерируемых в таких системах запросов пользователей является регламентными, а также тем, что даже произвольные запросы в таких системах могут иметь некоторую степень общности, поскольку они формируются в рамках решения регулярных управленческих задач.

Еще одним примером систем, в которых информационно-технологическое резервирование целесообразно, служат распределенные информационные системы многофилиальных банков, филиалы которых распределены по территориально удаленным городам Российской Федерации. Этот пример особенно актуален, так как все российские банки обязаны на регулярной основе готовить большое количество стандартных отчетных форм для представления в Банк России.

Информационно-технологический резерв размещается по узлам распределенной системы в виде некоторого числа идентичных копий. Распределенные по узлам сети копии ИТР, содержащие сформированные промежуточные массивы данных, используются при обработке типовых запросов пользователей, инициируемых после создания ИТР. Копии ИТР, размещенные по нескольким узлам сети, обеспечивают кратчайший путь и минимальное время доступа пользователей к необходимой им информации. Кроме того, наличие в узлах сети копий ИТР позволяет сократить время обработки запросов, поскольку для их обработки используются заранее подготовленные промежуточные массивы данных. Использование ИТР, таким образом, позволяет увеличить эффективность функционирования распределенной системы благодаря уменьшению времени реакции системы на запросы.

Кроме того, при регулярном создании в системе обработки данных полных копий (full backup) ИТР появляется возможность обработки запросов к историческим данным. Эти полные исторические копии можно использовать и для восстановления текущей копии ИТР в случае ее частичного или полного разрушения по разным причинам.

1.2. Цели создания информационно-технологического резерва

Конечные цели проектирования и создания информационно-технологического резервирования заключаются в:

- разработке множества процедур обработки запросов, предназначенных для предварительной обработки информации;
- формировании промежуточных массивов данных ИТР с помощью разработанных процедур;

— распределении ИТР (промежуточных массивов данных и множества процедур) по узлам РСОД централизованным или децентрализованным образом.

Проектирование ИТР осуществляется в предположении, что предметная область и множество запросов пользователей РСОД считаются заданными. Взаимодействие пользователей с РСОД представляется в виде множества формируемых ими типовых запросов. Каждый такой запрос можно отнести к некоторому определенному типу, обработка которого описывается графом исполнения.

Так как ИТР представляет собой составную часть РСОД, то его проектирование и создание ведется в рамках процесса проектирования и создания всей РСОД.

Процесс проектирования и создания ИТР в рамках РСОД состоит из трех этапов.

Этап 1. Анализ предметных областей пользователей. Проводится анализ предметных областей пользователей системы. Уточняются входные, промежуточные и итоговые данные. Уточняется перечень и последовательность выполнения процедур обработки этих данных при обработке запросов пользователей. Каждый запрос должен быть отнесен к тому или иному типу запросов. Каждый тип запросов характеризуется определенным набором используемых данных и процедур их обработки.

Для выполнения работ на данном этапе целесообразно воспользоваться получившими широкое распространение и неоднократно проверенными на практике формальными моделями и методами анализа предметных областей пользователей информационно-управляющих систем, разработанных в Институте проблем управления им. В.А. Трапезникова РАН [4, 5, 16–20]. Данные модели представляют собой совокупность взаимосвязанных матричных и графовых моделей и используются на этапе анализа предметных областей пользователей.

Результатом работ данного этапа являются множество различных типов входных, промежуточных и выходных данных и множество процедур обработки данных, используемых в процессе обработки запросов пользователей. Информационные связи между процедурами формализуются в форме мультиграфа, вершинами которого служат процедуры, а связывающие эти вершины дуги раскрашены общими для смежных процедур информационными элементами.

В графовой интерпретации задача анализа предметных областей пользователей состоит в разбиении полученного мультиграфа на подграфы с минимальным числом связей между ними [16, 17].

Этап 2. Синтез программного и информационного обеспечения информационно-технологического

резерва. На основе данных, полученных в результате выполнения работ первого этапа, проводится синтез информационного и программного обеспечения ИТР, т. е. определяются промежуточные и выходные данные и наборы процедур обработки данных, которые объединяются в программные модули. Синтез ИТР производится на основе принципа модульности, который позволяет оптимизировать состав и взаимосвязи отдельных компонентов программного и информационного обеспечения ИТР [16, 17, 20, 21].

В работе [21] дано содержательное определение ряда понятий, употребляемых для формализации и описания основных положений анализа и синтеза модульных систем обработки данных. Будем пользоваться следующими определениями.

Запрос — требование (заявка) пользователя на обработку данных, которое характеризуется некоторым составом обрабатываемых данных, результатом обработки данных и ограничением на время обработки.

Тип запросов — множество запросов, характеризующееся одинаковыми обрабатываемыми данными и набором процедур их обработки. К одному типу относятся такие запросы, для обработки которых используются одинаковый набор данных и одинаковое множество процедур их обработки, выполняемых системой в одинаковой последовательности.

Информационный элемент (ИЭ) — наименьшая, неделимая часть данных, допускающая независимое использование в процессе обработки данных.

Процедура обработки данных — формализованное действие по преобразованию ИЭ или группы ИЭ в другой ИЭ или группу ИЭ в соответствии с логикой обработки запросов.

Задача РСОД — набор процедур и обрабатываемых ими ИЭ, необходимых для обслуживания запросов одного или нескольких типов.

На основе результатов анализа предметных областей и регулярных запросов пользователей определяются характеристики ИТР:

- множество задач по обработке запросов;
- множество процедур обработки данных, используемых при решении задач по обработке запросов;
- вхождение процедур обработки данных в состав программных модулей;
- множество ИЭ, связанных с процедурами обработки данных и классифицируемых по типам: входной, промежуточный или выходной элемент данных;
- вхождение ИЭ в состав массивов данных;
- варианты взаимодействия процедур обработки данных с ИЭ;
- технологическая матрица смежности;



- множество допустимых последовательностей выполнения процедур;
- характеристики процедур и ИЭ.

Этап 3. Размещение ИТР по узлам сети. Единственная задача данного этапа состоит в оптимальном размещении копий ИТР по серверам РСОД. Цель такого размещения — повышение производительности, надежности и снижение стоимости функционирования распределенной системы. В качестве критериев оптимальности размещения копий ИТР могут использоваться такие критерии, как минимум среднего времени обработки запросов пользователей, минимум стоимости функционирования РСОД и максимум вероятности получения ответа на запрос пользователя.

Модели и методы анализа предметных областей пользователей подробно описаны в цитированной выше литературе, поэтому на них останавливаться не будем, а более подробно рассмотрим работы второго и третьего этапов процесса создания ИТР.

2. ФОРМУЛИРОВКА ЗАДАЧИ СИНТЕЗА ПРОГРАММНОГО И ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ИНФОРМАЦИОННО-ТЕХНОЛОГИЧЕСКОГО РЕЗЕРВА

Рассмотрим постановку задачи синтеза ИТР на этапе техно-рабочего проектирования распределенной системы обработки данных.

Введем обозначения:

$A = \{\alpha_1, \alpha_2, \dots, \alpha_r, \dots, \alpha_R\}$ — множество процедур обработки данных;

$D = \{d_1, d_2, \dots, d_r, \dots, d_L\}$ — множество ИЭ, обрабатываемых процедурами множества A ;

$W = \|\omega_{rl}\|$ — технологическая матрица смежности, которая отражает взаимосвязи ИЭ с процедурами их обработки, в которой:

$$\omega_{rl} = \begin{cases} +1, & \text{если } l\text{-й ИЭ — исходный} \\ & \text{для } i\text{-й процедуры,} \\ -1, & \text{если } l\text{-й ИЭ — результат} \\ & r\text{-й процедуры,} \\ 0, & \text{если } l\text{-й ИЭ не используется} \\ & r\text{-й процедурой;} \end{cases}$$

$W^c = \|\omega_{rl}^c\|$ и $W^3 = \|\omega_{rl}^3\|$ — соответственно матрицы взаимосвязей ИЭ с процедурами при считывании и записи, где

$$\omega_{rl}^{c(3)} = \begin{cases} 1, & \text{если } l\text{-й ИЭ считывается (записывается)} \\ & r\text{-й процедурой,} \\ 0 & \text{в противном случае;} \end{cases}$$

$\Omega(A)$ — множество допустимых последовательностей выполнения процедур при решении задачи обработки запросов, определенное на множестве процедур;

$w \in \Omega(A)$ — конкретная заданная последовательность выполнения процедур обработки данных.

В дополнение к множеству $\Omega(A)$ определим множество P возможных последовательностей выполнения процедур в виде матрицы $P = \|P_{r\theta}\|$, где

$$P_{r\theta} = \begin{cases} 1, & \text{если } r\text{-я процедура может быть выполнена} \\ & \theta\text{-й в заданной последовательности} \\ & \text{реализации процедур при обработке данных,} \\ 0 & \text{в противном случае;} \end{cases}$$

τ_v — среднее время считывания v -го модуля из внешней памяти в оперативную память сервера;

t_f^c — среднее время считывания f -го массива данных из внешней памяти в оперативную память сервера;

t_f^3 — среднее время записи результатов обработки в f -й массив данных.

Методы, которые позволяют получить оценки значений величин τ_v , t_f^c и t_f^3 рассмотрены в работах [18, 19, 22].

В общем случае компьютерная сеть неоднородна и строится на базе серверов и каналов связи с различными техническими характеристиками, что существенно увеличивает вычислительную сложность задачи. Поэтому с целью сокращения размерности задачи будем пользоваться для данных величин усредненными оценками на множестве всех серверов и каналов связи.

Введем переменные:

$$y_{vl}^{c(3)} = \begin{cases} 1, & \text{если } \sum_{r=1}^R \omega_{rl}^{c(3)} x_{rv} \geq 1, \\ 0, & \text{если } \sum_{r=1}^R \omega_{rl}^{c(3)} x_{rv} < 1; \end{cases}$$

$$y_{vf}^{c(3)} = \begin{cases} 1, & \text{если } \sum_{l=1}^L y_{vl}^{c(3)} z_{lf} \geq 1, \\ 0, & \text{если } \sum_{l=1}^L y_{vl}^{c(3)} z_{lf} < 1; \end{cases}$$

$$x_{rv} = \begin{cases} 1, \text{ если } r\text{-я по порядку выполнения процедура} \\ \text{включается в состав } v\text{-го модуля,} \\ 0 \text{ в противном случае;} \end{cases}$$

$$z_{vf} = \begin{cases} 1, \text{ если первый ИЭ включается в } f\text{-й массив,} \\ f = \overline{1, F}; \quad F \leq L. \\ 0 \text{ в противном случае.} \end{cases}$$

Переменные $y_{vl}^{c(3)}$ и $y_{vf}^{c(3)}$ служат для формализации взаимосвязи модулей ИТР с ИЭ и массивами данных при считывании (записи) в процессе обмена с внешней памятью серверов.

Формулировка задачи. Введенные обозначения и переменные позволяют сформулировать задачи поиска оптимальной структуры ИТР для заданной предметной области и множества типовых запросов пользователей РСОД [16, 21].

На этапе проектирования модульных систем обработки данных принимаются следующие основные критерии [16]:

- минимум времени, затраченного на обмен данными между оперативной памятью и внешними запоминающими устройствами при решении задач;
- минимум сложности межмодульного интерфейса;
- минимум объема неиспользуемых данных при пересылках между оперативной и внешней памятью серверов системы;
- минимум технологической сложности алгоритмов обработки данных;
- максимум производительности системы при решении задач и др.

Рассмотрим формулировку задачи синтеза оптимальной структуры программного и информационного обеспечения ИТР. В качестве критерия оптимизации примем минимальное общее время обмена с внешней памятью серверов при обработке запросов пользователей. Данный критерий имеет большое значение при синтезе программного и информационного обеспечения систем обработки данных, так как время, затрачиваемое на обмен данными, во многом определяет быстродействие системы.

Задача формулируется таким образом: найти

$$\min_{(x_{rv}, z_{vf})} \left\{ \sum_{r=1}^R \sum_{v=1}^V x_{rv} (1 - x_{r+1, v}) \times \left[\tau_v + \sum_{f=1}^F (z_{vf}^c t_f^c + z_{vf}^3 t_f^3) \right] \right\} \quad (1)$$

при ограничениях на:

- общее число процедур обработки данных в составе каждого модуля:

$$\sum_{r=1}^R x_{rv} \leq \bar{M}, \quad v = 1, 2, \dots, V, \quad (2)$$

где $\bar{M}$ — максимально допустимое число процедур в модуле;

- число ИЭ, обрабатываемых процедурами каждого модуля,

$$\sum_{l=1}^L y_{vl} \leq \bar{L}, \quad v = 1, 2, \dots, V, \quad (3)$$

здесь $\bar{L}$ — максимально допустимое число ИЭ, обрабатываемых v -м модулем;

- сложность межмодульного взаимодействия в ИТР

$$\sum_{l=1}^L \sum_{v=1}^{V-1} \sum_{v'=v+1}^V y_{vl} y_{v'l} \leq S, \quad (4)$$

где S — максимально допустимая мощность информационного интерфейса между модулями ИТР, т. е. максимальное допустимое число переменных, общих для программных модулей ИТР;

- сложность информационного интерфейса между отдельными парами модулей

$$\sum_{l=1}^L y_{vl} y_{v'l} \leq S_{vv'}, \quad (5)$$

где для заданных модулей v и v' величина $S_{vv'}$ — это максимально допустимое число общих ИЭ, обрабатываемых модулями v и v' ;

- однократность включения процедур в программные модули ИТР

$$\sum_{v=1}^V x_{rv} = 1, \quad r = 1, 2, \dots, R; \quad (6)$$

- включение отдельных процедур в состав одного модуля

$$x_{rv} + x_{r'v} \leq 1 \quad (7)$$

для заданных процедур α_r и $\alpha_{r'}$, $v = 1, 2, \dots, V$;



— передачу управления из модуля до завершения обработки данных всеми процедурами этого модуля

$$\sum_{r=1}^R x_{rv}(1 - x_{r+1,v}) = 1, \quad v = 1, 2, \dots, V; \quad (8)$$

— дублирование ИЭ в массивах данных

$$\sum_{f=1}^F z_{vf} = k, \quad k = 1, 2, \dots, K; \quad (9)$$

где K — допустимая степень дублирования ИЭ в массивах данных;

— размер записи каждого массива данных

$$\sum_{l=1}^L z_{vf} \leq M, \quad f = 1, 2, \dots, F, \quad (10)$$

где N — максимально допустимое число ИЭ в массиве.

Если существуют трудности при оценке временных характеристик работы серверов компьютерной сети с программными модулями и массивами данных, то в этом случае в качестве критерия оптимизации следует принять минимальное число обращений к внешней памяти при обработке запросов пользователей.

Данная задача является вариантом задачи (1)—(10) и формулируется таким образом: найти

$$\min_{(x_{rv}, z_{vf})} \left\{ \sum_{r=1}^R \sum_{v=1}^V x_{rv}(1 - x_{r+1,v}) \times \left[1 + \sum_{f=1}^F (z_{vf}^c + z_{vf}^3) \right] \right\} \quad (11)$$

при ограничениях (2)—(10).

Рассмотренные задачи (1)—(11) синтеза модульной структуры программного и информационного обеспечения ИТР являются нелинейными задачами целочисленного программирования комбинаторного типа, для решения которых целесообразно использовать алгоритмы, приведенные в работе [21].

ВЫВОДЫ

Рассмотренный процесс проектирования и создания ИТР в РСОД позволяет найти эффективное решение задачи использования ИТР в распределенных системах, в которых частоты возникновения запросов определенного типа относительно равномерно распределены по узлам компьютерной сети. Другими словами, когда в сети отсутствуют

сегменты, в которых наблюдается явное превышение интенсивности запросов одного или нескольких типов над интенсивностью запросов аналогичного типа, возникающих в других сегментах сети. Такая ситуация может иметь место в работе распределенной информационной системы многофилиального банка, в котором реализована технология обслуживания клиента в любом иногороднем филиале. В этом случае пользователи отделений филиала NN банка в городе N в большинстве случаев будут работать с данными клиентов банка, зарегистрированных в отделениях филиала этого же города N. И только в случае, когда в филиал NN в городе N приходит клиент, живущий и зарегистрированный в филиале XX города X, запросы к данным об этом клиенте из филиала NN будут направлены в сегмент сети, обслуживающий филиал XX города X.

В подобной ситуации процесс проектирования и синтеза ИТР усложняется. В этом случае работы на этапах синтеза и размещения ИТР распараллеливаются на несколько потоков. Каждый такой поток работ предназначен для синтеза и последующего размещения в сети ИТР, предназначенного для обработки одного или нескольких типов запросов. В данном потоке работ будут рассматриваться сегменты сети, в которых интенсивность запросов этих типов существенно выше, чем в других сегментах. В данной статье такая постановка задачи не рассматривается и служит предметом дальнейших исследований.

Отметим, что при проектировании ИТР для крупномасштабных РСОД с большим и очень большим количеством пользователей потребуются выполнить и существенный объем работ по анализу предметных областей пользователей системы:

- территориальная распределенность пользователей;
- перечень запросов и разбиение их по различным типам;
- запрашиваемая пользователями информация;
- частота возникновения запросов в узлах сети;
- выделение групп пользователей, запрашивающих аналогичные данные.

Очевидно, что этот процесс потребует выделения значительных ресурсов бизнес- и системных аналитиков, а также архитекторов РСОД.

В настоящей статье остался открытым вопрос о том, как необходимо разместить по узлам компьютерной сети спроектированный ИТР, чтобы достигался наибольший эффект от его использования в РСОД. Задача оптимального размещения копий ИТР по узлам РСОД будет рассмотрена в следующей статье.

ЗАКЛЮЧЕНИЕ

Рассмотренный новый подход к использованию информационной и программной избыточности в виде распределенных копий ИТР обеспечивает повышение реакции РСОД при обработке типовых, регулярных запросов пользователей таких систем. Множество копий ИТР после их создания и размещения по узлам сети служит информационной базой для РСОД при обработке запросов пользователей в будущем.

Использование такого вида резервирования, созданного для обработки типовых запросов, которые будут сформированы в будущем, позволяет ускорить процесс их обработки. Это ускорение достигается благодаря тому, что промежуточные данные, необходимые для этой обработки, подготавливаются заранее, до поступления самих запросов.

Описаны основные этапы формирования ИТР, перечислены преимущества этого вида резервирования и выделены важные вопросы, на которые необходимо уделить особое внимание при проектировании РСОД, использующих ИТР. Сформулированы задачи поиска оптимальной структуры ИТР с учетом предметных областей и множества типовых запросов пользователей распределенной системы.

Практика использования в РСОД информационно-технологического резервирования показывает, что применение такого вида резервирования приводит к экономии на 10–20 % затрат на обработку типовых запросов пользователей.

ЛИТЕРАТУРА

1. Таненбаум Э., ван Стеен М. Распределенные системы. Принципы и парадигмы. — СПб.: Питер, 2003. — 877 с.
2. Ozsu M.T., Valduriez P. Principles of distributed database systems. — 2nd ed. — NJ, USA, Upper Saddle River: Prentice-Hall, Inc., 1999. — 845 p.
3. Коннолли Т., Бег К. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. — 3-е изд. — М.: Изд. дом «Вильямс», 2003. — 1440 с.
4. Кульба В.В., Ковалевский С.С., Косяченко С.А., Сиротюк В.О. Теоретические основы проектирования оптимальных структур распределенных баз данных. — М.: СИНТЕГ, 1999. — 660 с.
5. Сиротюк В.О. Разработка и исследование моделей и методов анализа и синтеза оптимальных структур баз данных иерархического и сетевого типов: дис. ... канд. техн. наук. — М.: ИПУ РАН, 1984. — 270 с.
6. Карсанидзе Т.В. Резервирование, восстановление и регистрация информации в автоматизированных системах управления, функционирующих на базе локальных сетей ЭВМ: автореф. дис. ... канд. техн. наук. — М.: ИПУ РАН, 1992. — 23 с.

7. Mansouri N. Adaptive data replication strategy in cloud computing for performance improvement // *Frontiers of Computer Science*. — 2016. — Vol. 10, iss. 5. — P. 925–935.
8. Sahoo J., Salahuddin M.A., Glietho R. A Survey on Replica Server Placement Algorithms for Content Delivery Networks // *IEEE Communications Surveys & Tutorials*. — 2016. — 30 p. [Электронный ресурс]. — URL: <https://arxiv.org/ftp/arxiv/papers/1611/1611.01729.pdf> (дата обращения: 10.03.2018).
9. Soury A., Rahmani A.M. A Survey for Replica Placement Techniques in Data Grid Environment // *International Journal of Modern Education and Computer Science*. — 2014. — Iss. 5. — P. 46–51.
10. Abdalla H.I. A synchronized design technique for efficient data distribution // *Computers in Human Behavior*. — 2014. — Vol. 30. — P. 427–435.
11. Singh A., Kahlon K.S. Non-replicated dynamic data allocation in distributed database system // *International Journal of Computer Science and Network Security*. 2009. — Vol. 9, N 9. — P. 176–180.
12. Чернышев Г.А. Обзор подходов к организации физического уровня в СУБД // *Тр. СПИИРАН*. — 2013. — Вып. 1 (24). — С. 222–275.
13. Шавелев Л.В. Оперативная аналитическая обработка данных: концепции и технологии. — URL: http://www.olap.ru/basic/olap_and_ida.asp (дата обращения: 6.07.2017).
14. Туманов В. Data Warehouse: с чего начать? // *PC Week*. — 1998. — № 29 (153). — URL: <https://www.itweek.ru/infrastructure/article/detail.php?ID=48156> (дата обращения: 6.07.2017).
15. Standen J. Data Warehouse vs Data Mart. — URL: <http://www.datamartist.com/data-warehouse-vs-data-mart>. (дата обращения: 6.07.2017).
16. Кузнецов Н.А., Кульба В.В., Ковалевский С.С. Методы анализа и синтеза модульных информационно-управляющих систем. — М.: Физматлит, 2002. — 800 с.
17. Казиев Г.З., Кузнецов Н.А., Кульба В.В. и др. Модели, методы и средства анализа и синтеза модульных информационно-управляющих систем // *Автоматика и телемеханика*. — 1993. — № 6. — С. 3–59.
18. Мамиконов А.Г., Кульба В.В. Синтез оптимальных модульных СОД. — М.: Наука, 1989. — 276 с.
19. Мамиконов А.Г., Кульба В.В., Косяченко С.А. Типизация разработки модульных систем обработки данных. — М.: Наука, 1989. — 163 с.
20. Кульба В.В., Мамиконов А.Г. Методы анализа и синтеза оптимальных модульных систем обработки данных (Обзор) // *Автоматика и телемеханика*. — 1980. — № 11. — С. 152–170.
21. Информационное обеспечение систем организационного управления (теоретические основы): в 3-х ч. — Ч. 2. Методы анализа и проектирования информационных систем / под ред. Е.А. Микрина, В.В. Кульбы. — М.: Изд-во физ.-мат. лит., 2011. — 496 с.
22. Ашимов А.А., Мамиконов А.Г., Кульба В.В. Оптимальные модульные системы обработки данных. — Алма-Ата: Наука, 1981. — 186 с.

Статья представлена к публикации членом редколлегии В.В. Кульбой.

Сомов Сергей Константинович — канд. техн. наук, вед. инженер, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ ssomov2016@ipu.ru.

УПРАВЛЕНИЕ АРХИТЕКТУРОЙ КИБЕРБЕЗОПАСНОСТИ АСУТП АТОМНЫХ ЭЛЕКТРОСТАНЦИЙ

Д.И. Бабаев, А.Г. Полетыкин, В.Г. Промыслов, М.Ю. Тимофеев

Рассмотрены методы оценки риска, назначения уровней кибербезопасности, требования к мерам защиты и построения архитектуры кибербезопасности. Сформулированы основные задачи обеспечения кибербезопасности. Обозначены принципы построения архитектуры кибербезопасности. Обсужден пример практической реализации архитектуры.

Ключевые слова: информация, кибербезопасность, безопасность, автоматизированная система управления, атомная электростанция.

ВВЕДЕНИЕ

Безопасность общества и государства в существенной мере зависит от безопасности системообразующих инфраструктурных объектов, к которым относятся, например, объекты транспортной, химической, энергетической и других сфер.

На различных уровнях управления обращаются огромные массивы информации, требующие соответствующей защиты от несанкционированных воздействий, таких как, например, компьютерные атаки. Наличие цифровых систем управления, непосредственно связанных с технологическим процессом и выполняющих также противоаварийные функции, расширяет возможное пространство атаки, поскольку часть данных в системе может быть критически важным активом для поддержания функционирования объекта. Изложенное справедливо и для системы управления технологическими процессами на атомной электростанции (АЭС). На АЭС цифровая АСУТП выполняет как функции безопасности, так и функции нормальной эксплуатации [1]. Проблема обеспечения безопасности АСУТП АЭС комплексна и связана с обеспечением промышленной (энергетической, ядерной и радиационной, функциональной) безопасности [2, 3], а из-за высокой степени централизации управления на основе компьютеризированных систем на объекте, и с обеспечением информационной и кибербезопасности [4]. Проблематика обеспечения кибербезопасности критических систем описана в ряде работ и документов российских и международных организаций. Например, описаны общие подходы к обеспечению кибербезопасности [5], представлен подход к оценке кибербезопасности АСУТП на базе методов структурно-логи-

ческого анализа надежности и безопасности с применением вероятностных моделей [6], рассмотрен метод анализа кибербезопасности на основе теории графов и дискреционной модели доступа [7]. Среди комплексных работ отметим профильные документы, разработанные Международной электротехнической комиссией (МЭК) и МАГАТЭ [8–10], серию стандартов ISO/IEC 27000, документ [11] в части общих принципов обеспечения безопасности цифровых систем управления и др. Однако ни в специальных, ни публикациях общего характера на данный момент не удалось полностью учесть всю специфику и сложность АСУТП АЭС.

В настоящей статье рассмотрена общая модель взаимодействия различных компонентов безопасности АЭС, стратегия управления кибербезопасностью в рамках жизненного цикла АСУТП АЭС. Систематизирован процесс анализа рисков, которые позволяют оценивать кибербезопасность проектных решений в части архитектуры безопасности и определяют функционирование компонентов АСУТП АЭС. Рассмотрена типовая реализация архитектуры кибербезопасности АСУТП.

1. ОБЩАЯ СТРАТЕГИЯ УПРАВЛЕНИЯ КИБЕРБЕЗОПАСНОСТЬЮ

В международной практике для выделения кибербезопасности из информационной безопасности принята модель МАГАТЭ [12], которая в основном используется в настоящее время и на национальном уровне РФ для атомной промышленности¹.

¹ С 1 января 2018 г. общая модель взаимодействия на национальном уровне определяется федеральным законом [5].

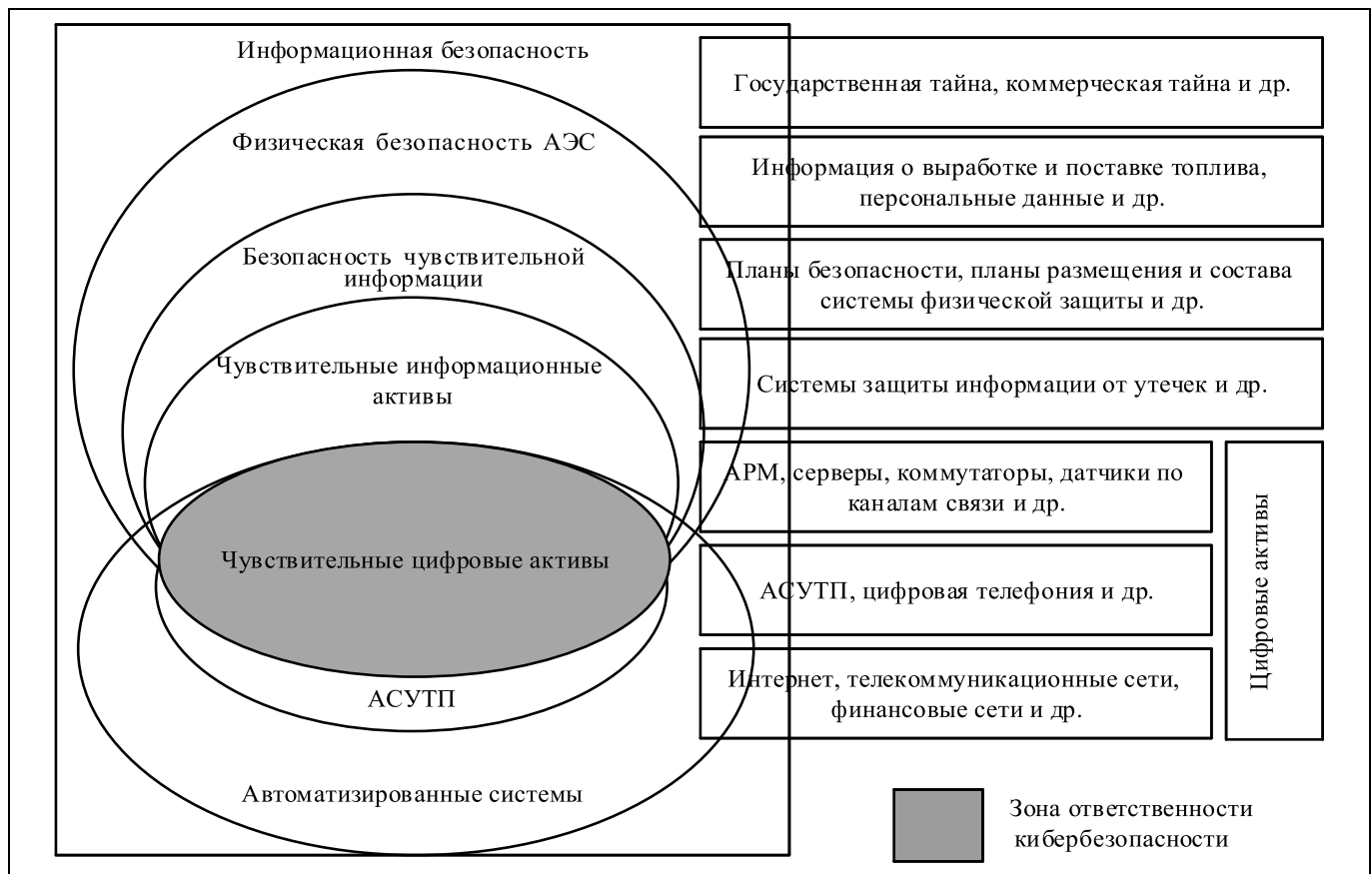


Рис. 1. Общая модель взаимодействия информационной и кибербезопасности для атомной промышленности

В соответствии с этой моделью взаимодействия (рис. 1) в рамках обеспечения информационной безопасности выделяются следующая ее внутренняя структура и соответствующие процессы:

- данные классифицируются по уровням важности и к ним предъявляются общие требования по защите информации [13]; требования по защите информации распространяются как на собственно данные в электронной форме, так и на компьютерные системы, используемые для их хранения, передачи и обработки;

- дополнительные требования на информацию налагаются в рамках законов и регулирующих документов, связанных с атомной энергетикой [2];

- собственник АЭС, в соответствии с требованиями в рамках программы информационной безопасности, определяет политики и процедуры безопасности для критически важной информации, нарушение доступности, целостности или конфиденциальности которой может привести к нарушению штатного режима функционирования АСУ [11];

- собственником идентифицируются и выделяются критически важные информационные активы;

- собственник идентифицирует цифровой актив и тип актива, который осуществляет обработку, хранение, передачу информации в цифровой форме.

Кибербезопасность заключается в поддержании в заданных пределах значений рисков (экономических, экологических, социальных), связанных с возможными нарушениями (умышленными и неумышленными) доступности, целостности или конфиденциальности информации (алгоритмов, данных и сигналов) в АСУТП АЭС.

Процесс управления кибербезопасностью АСУТП АЭС в общем случае можно представить в виде последовательности выполнения следующих шагов [8, 9]:

- по оценкам риска на системном уровне назначаются уровни кибербезопасности для подсистем АСУТП;

- в соответствии с уровнем кибербезопасности назначаются требования к мерам защиты;

- определяется архитектура кибербезопасности АСУТП, в том числе правила разграничения доступа;

- оценивается остаточный риск на уровне подсистем и их элементов и при необходимости на-



значаются дополнительные меры кибербезопасности.

Кибербезопасность АСУТП АЭС достигается посредством комплекса организационно-технических мероприятий. Весь комплекс мероприятий направлен на выявление (обнаружение), различного вида угроз, на защиту (реагирование) от этих угроз и устранение последствий от реализации этих угроз. Реализация таких угроз способна привести к нарушению целостности, доступности или конфиденциальности обрабатываемой и утрате управляемости, наблюдаемости и устойчивости технологического процесса АЭС. Кибербезопасность АСУТП АЭС обеспечивается в рамках соответствующей программы, которая определяет и координирует деятельность организаций, вовлеченных в жизненный цикл АСУТП (разработчики, проектировщики, поставщики и др.) по выявлению и ликвидации (нейтрализации) угроз объектам АСУТП, снижению рисков и размера возможного ущерба на стадиях жизненного цикла АСУТП.

Основополагающим документом по кибербезопасности АСУТП АЭС служит политика безопасности, разрабатываемая как для каждого из этапов жизненного цикла, так и для каждого из уровней представления системы. Политики безопасности определяются на различных уровнях, начиная от политики управления или администрирования, устанавливаемой на уровне предприятия, и заканчивая политикой эксплуатации, описывающей конкретные меры по управлению безопасностью. Дополнением к политике безопасности служат процедуры и регламенты. Разделение политики (описания) и действий (процедур) по воплощению политики безопасности позволяет более гибко реагировать на изменение рисков, связанных с кибербезопасностью. Поскольку угрозы реализуются посредством уязвимостей в политике безопасности, для уменьшения связанного с кибербезопасностью риска возможно воздействовать как на угрозы, так и на уязвимости. Уязвимости, в отличие от угроз, подвержены прямому воздействию со стороны владельца защищаемого объекта, поэтому управление риском посредством минимизации уязвимостей является более гибким и чаще применяемым методом.

Основные задачи обеспечения кибербезопасности должны заключаться в:

- плановом выявлении и устранении угроз кибербезопасности и их источников;
- развитию и совершенствовании системы обеспечения кибербезопасности, включая разработку новых и совершенствование существующих способов, методов и средств выявления, оценки, прогнозирования, нейтрализации и ликвидации угроз, а также средств и методов противодействия этим угрозам;

— эффективном противодействии угрозам кибербезопасности;

— выявлении, предотвращении и ликвидации последствий воздействий, вызывающих нарушения целостности, доступности и конфиденциальности информации и данных, приводящих к нарушению технологического процесса;

— сертификации средств защиты информации, программного обеспечения, технологий, их разработки и применения в соответствии с требованиями безопасности;

— совершенствовании приемов, способов, методов и средств защиты от кибератак;

— своевременной и непрерывной подготовке кадров.

2. ОЦЕНКА РИСКА НА СИСТЕМНОМ УРОВНЕ

Контекст кибербезопасности, как на общесистемном уровне, так и на уровне элементов подсистем АСУТП АЭС основан на концепциях угроз, рисков и контрмер, а также взаимосвязях между ними. Анализ рисков должен проводиться для взаимосвязанной системы «человек—машина». Для АСУТП АЭС и информационных систем используется простая контекстная модель анализа рисков ISO [14] (рис. 2). Специфичным для АСУТП АЭС является контекст, связанный с необходимостью обязательного рассмотрения сценариев инцидентов ядерной и радиационной безопасности (ЯРБ) из-за очень высокого уровня риска, связанного с этими сценариями и намного превосходящего любой другой возможный риск. Реализуемые меры кибербезопасности позволяют обеспечить надежное и качественное функционирование АЭС и уменьшение рисков нарушения ЯРБ при эксплуатации уязвимостей цифровых компонентов АСУТП АЭС.

Стандартный подход к оценке рисков основан на сценарии ISO 27005 [15]. Сценарии риска строятся с учетом трех элементов.

- Активы. Первый шаг в классификации и обеспечении информации заключается в идентификации активов, важных для кибербезопасности, с учетом выбранного уровня представления системы. Для идентификации активов выполняется анализ функциональной безопасности с целью выделения активов, критически важных для обеспечения выполнения проектных функций всей системы.
- Уязвимость, которая связана с воздействием на актив в обход политики безопасности системы.
- Угроза, которая связана с вероятностью (частотой) использования уязвимости.

В таком понимании контекста кибербезопасности предлагается выделять следующие риски:

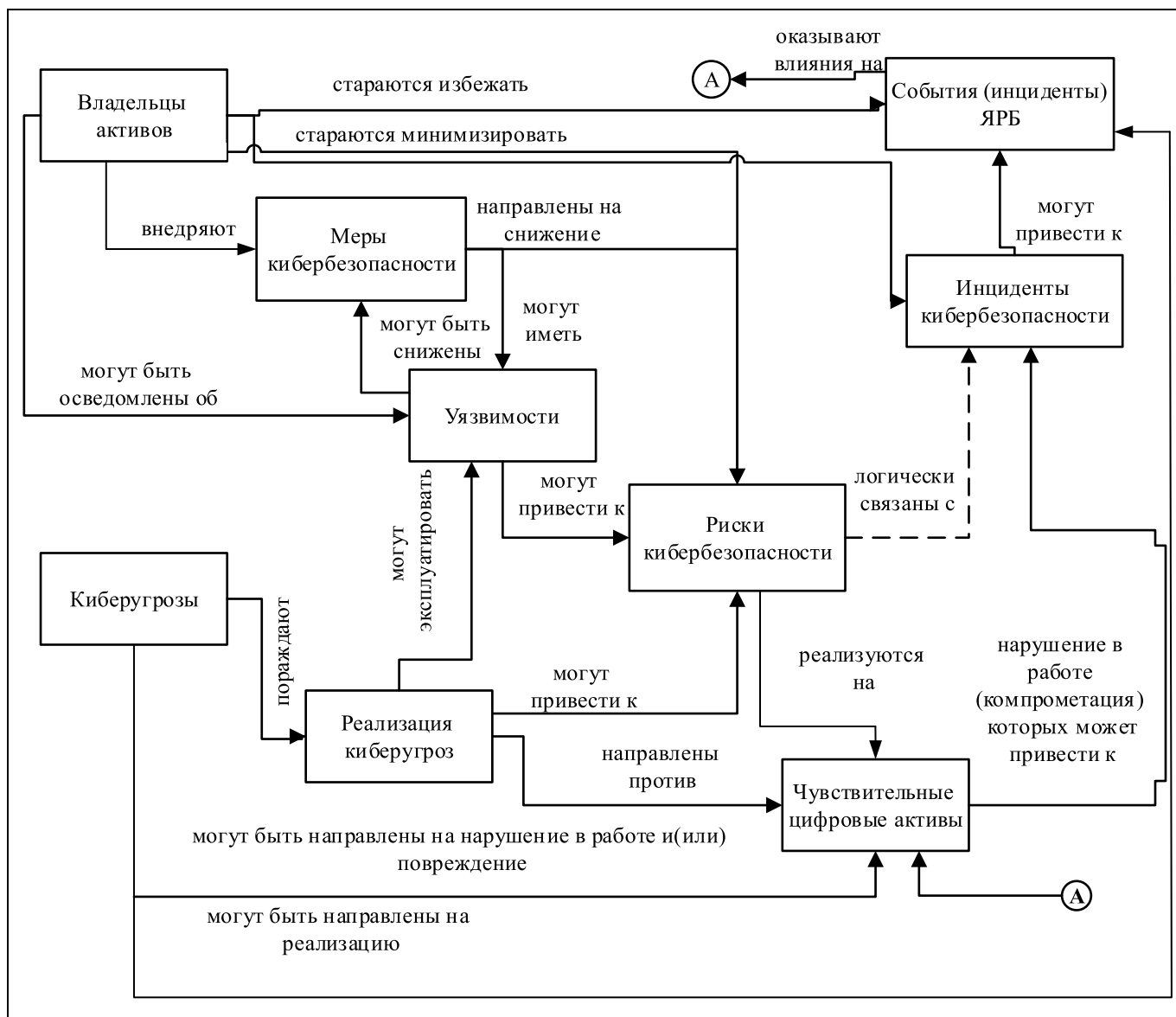


Рис. 2. Взаимосвязи между элементами контекста анализа риска

- системный, определяемый системообразующими факторами;
- правовой, определяемый оценками действующих юридических норм;
- экономический, определяемый эффективностью функционирования объекта;
- функциональный, определяемый факторами, влияющими на достижение цели, заложенной при проектировании компонента.

Многие руководства и нормативные документы, например, ISO 27005, определяют стандартный набор уязвимостей для информационных систем, а также вводят градации для активов, уязвимостей и угроз вместе с правилами расчета риска для их комбинаций. Однако, как показывает практика,

для индустриальных систем из-за их узкой специализации стандартные подходы малоприменимы. Разработчики чаще всего вынуждены создавать собственные методики оценки риска, адаптированные к их системам.

Основная тенденция, которая прослеживается в общих подходах МЭК и МАГАТЭ, заключается в адаптации результатов оценки ядерной безопасности на системном и функциональном уровнях для классификации систем по кибербезопасности вместо проведения отдельной оценки риска вследствие нарушения целостности, доступности и конфиденциальности информации. Особенно такой подход часто используется на начальном этапе для общей оценки риска в системе. Схема назначения

уровней кибербезопасности в соответствии с категориями ядерной безопасности для выполняемых системой функций, применяемая МЭК [8], рассматривается далее.

3. НАЗНАЧЕНИЕ УРОВНЕЙ КИБЕРБЕЗОПАСНОСТИ

3.1. Непрямая оценка риска кибербезопасности

Для атомной отрасли характерно наличие глубокой проработки технологии и результатов по оценке ядерных рисков. Результаты представлены в виде классификаторов по ядерной опасности как для систем, так и для выполняемых системой функций. В атомной промышленности как на национальном, так и на международном уровне разработана система классификации, связывающая системы или функции с их влиянием на ядерную безопасность [16, 17]. Класс ядерной безопасности, присвоенный системе, в значительной мере определяет ценность системы, так как с ним однозначно связан ущерб в случае выхода системы из строя (шкала ущербов INES).

Рассмотрим подробнее схему классификации активов по уровням кибербезопасности, исходя из назначенных им классов по ядерной безопасности. В классификационной схеме МЭК определяется три уровня безопасности: S1, S2 и S3. При назначении уровней безопасности функции и системы должны анализироваться для определения максимальных последствий для безопасности АЭС в случае выполнения вредоносных действий или возникновения связанных с ними событий (рис. 3). Системы должны быть отнесены к уровням от S1 (наиболее строгие требования по кибербезопасности)

до S3 в соответствии с максимальными последствиями кибератаки для ядерной безопасности (ЯРБ) и экономических характеристик объекта. Алгоритм присвоения уровней кибербезопасности для АСУТП АЭС основан на категории функций безопасности, которые выполняет система:

- системы, выполняющие функции категории А, относятся к уровню кибербезопасности S1;
- системы, выполняющие функции категории В, а также системы с прямым (непосредственным) воздействием на безопасность АЭС или отказы которых могут привести к нарушению технологического процесса и введению в действие функций категории А, относятся к уровню кибербезопасности не ниже S2;
- системы, поддерживающие функции категории С в соответствии с максимально возможными последствиями от реализации кибератак на них, а также системы, обеспечивающие эксплуатацию и техническое обслуживание АЭС до уровня кибербезопасности S3.

При этом уровень кибербезопасности может быть повышен по итогам проведения углубленной оценки рисков.

Отметим, что МАГАТЭ приводит только классификацию по ЯРБ систем и оборудования в составе АСУТП АЭС [10], а также определяет число и свойства уровней кибербезопасности. Всего определяются пять уровней кибербезопасности, три высших из которых соответствуют уровням МЭК.

3.2. Прямая оценка риска кибербезопасности

Кроме классификации активов по кибербезопасности, с учетом их классификации по ЯРБ, может быть использована схема классификации,

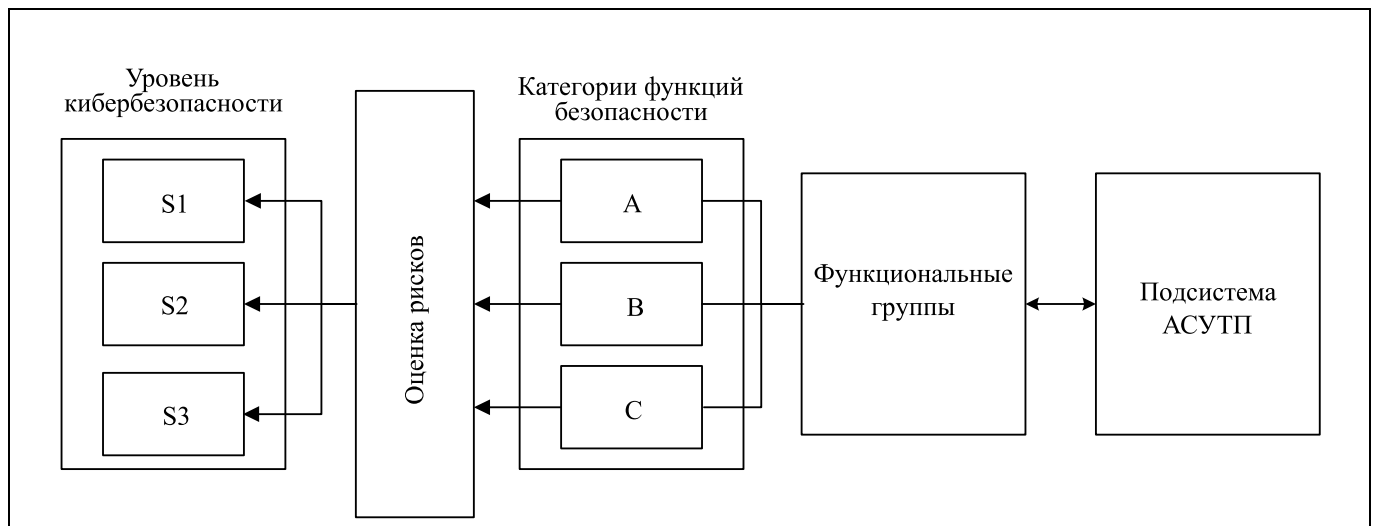


Рис. 3. Схема назначения уровней кибербезопасности

напрямую не связанная с ядерной безопасностью, например, схема приведенная в документе [11]. В ней также выделяются три уровня защищенности: К1 (самый высокий), К2 и К3. Для назначения класса защищенности определяется уровень значимости обрабатываемой информации в зависимости от степени возможного ущерба от нарушения целостности, доступности или конфиденциальности (при необходимости) обрабатываемой информации. Уровень значимости (критичности) информации определяется степенью возможного ущерба от нарушения ее целостности (неправомерное уничтожение или модифицирование), доступности (неправомерное блокирование) или конфиденциальности (неправомерные доступ, копирование, предоставление или распространение), в результате которых возможно нарушение штатного режима функционирования АСУ или незаконное вмешательство в процессы функционирования АСУ. Уровень значимости отдельных подсистем в составе АСУ определяется экспертным путем.

3.3. Сравнительный анализ методик

По мнению авторов правильное и квалифицированное применение перечисленных выше методик классификации активов (систем) АСУТП АЭС по кибербезопасности приводит к похожим результатам и реализации фактически аналогичного набора мер защиты.

На практике применение и выбор методик, в основном, определялось тем, на какой рынок — национальный или международный ориентирована защищаемая система, а также опытом работы разработчика. Для АСУТП АЭС, поставляемых на международный рынок, применялась обычно методика МЭК или МАГАТЭ, для отечественного рынка — методика, изложенная в [11]. Некоторое преимущество имеет методика классификации МЭК, специально ориентированная на применение в атомной промышленности — с простым и понятным алгоритмом, связанным с ЯРБ, что уменьшает объем работ по оценке риска кибербезопасности на начальном этапе.

4. ВЫБОР МЕР ПО ОБЕСПЕЧЕНИЮ КИБЕРБЕЗОПАСНОСТИ

Обеспечение кибербезопасности является составной частью работ по созданию и эксплуатации АСУТП АЭС. Обеспечение кибербезопасности осуществляется на стадиях (этапах) жизненного цикла АСУТП АЭС путем выполнения организационно-технических мероприятий, направленных на блокирование (нейтрализацию) угроз. Мероп-

риятия по обеспечению кибербезопасности должны быть соизмеримы по своей эффективности.

Если определенная мера кибербезопасности не может быть реализована по техническим, экономическим или иным причинам или если ее применение противоречит требованиям ЯРБ, то применяются некоторые компенсирующие меры (организационные, меры физической защиты и др.). Формальное обоснование эффективности применения мер должно быть продемонстрировано анализом угрозы и риска. Если компенсирующая мера, также не может быть применена, риск кибербезопасности, связанный с неприменением данной меры, должен быть принят.

Цель компенсирующих мер безопасности заключается в том, чтобы снизить уровень риска кибербезопасности до уровня, который будет поддерживать рекомендованный контроль безопасности.

Различные меры кибербезопасности по-разному воздействуют на целостность, доступность и конфиденциальность критических активов АСУТП АЭС. Применение мер кибербезопасности специфично для различных этапов жизненного цикла АСУТП АЭС.

Для обеспечения кибербезопасности АСУТП АЭС необходимы следующие организационно-технические мероприятия [5]:

- по определению объектов защиты;
- по формированию требований к защите от киберугроз;
- по классификации систем по кибербезопасности;
- по выделению актуальных угроз;
- по выбору мер обеспечения кибербезопасности;
- по проектированию и разработке системы обеспечения кибербезопасности;
- организационные мероприятия по обеспечению кибербезопасности;
- оценка соответствия требованиям кибербезопасности подсистем и АСУТП АЭС в целом;
- обеспечение кибербезопасности при вводе в эксплуатацию;
- обеспечение кибербезопасности в ходе эксплуатации;
- обеспечение кибербезопасности при выводе из эксплуатации.

В атомной промышленности такой список мер приводится в документах МЭК и МАГАТЭ [9, 10]. Для АСУТП других систем подобный список мер, отнесенный к уровням кибербезопасности, приведен в документе [11]. Набор мер МАГАТЭ и [11], в отличие от документов МЭК, не разделен по этапам жизненного цикла, что является недостатком.



5. АРХИТЕКТУРА КИБЕРБЕЗОПАСНОСТИ

5.1. Аспекты архитектуры кибербезопасности

Достигнутая реальная кибербезопасность АСУТП АЭС, как объекта защиты, на этапе эксплуатации в большой мере зависит от той архитектуры кибербезопасности [18], которая была реализована в ходе выполнения программы кибербезопасности. Для разработки, внедрения и обслуживания программы кибербезопасности предполагается реализация следующих действий:

- анализ цифровых систем и локально-вычислительных сетей ядерных установок;
- выявление и оценка критических ресурсов в смысле безопасности;
- внедрение безопасной архитектуры в соответствии с определенными нормативами;
- анализ потенциальных рисков нарушений кибербезопасности;
- разработка детальной архитектуры кибербезопасности.

Архитектура кибербезопасности определяет связанную с кибербезопасностью структуру АСУТП АЭС как системы систем, включая основные функции, класс и границы каждой системы, взаимосвязь или независимость систем, приоритетность целей, одновременно действующих в системе, и порядок взаимодействия между человеком и системой. Архитектура кибербезопасности также включает в себя план и набор правил, описывающих сервисы безопасности, которые должны обеспечивать защиту системы и ее пользователей, необходимые меры для реализации этих сервисов, требуемые показатели эффективности функционирования элементов.

Архитектура кибербезопасности охватывает такие аспекты как:

- рекомендуемые компоновки сетей передачи данных между элементами системы;
- рекомендуемые конфигурации межсетевого экрана;
- процедуры авторизации и аутентификации пользователей;
- барьеры между различными сетями управления процессами;
- применение беспроводных коммуникаций;
- управление обновлениями и патчами (включая аутентификацию при их проведении);
- управление средствами обеспечения кибербезопасности;
- усиление защиты систем путем закрытия программных портов, блокировки или отказа от неиспользуемых или опасных сервисов и отказа от применения съемных накопителей данных;
- доступ к внешним сетям передачи данных за пределами АСУТП АЭС (включая Интернет);

— надлежащее применение средств электронной связи между персоналом (например, электронной почты).

Архитектуры кибербезопасности АЭС строятся на основе системного подхода, при котором верхнеуровневые стратегические цели и принципы наследуются всеми функциями и активами объекта (включая персонал). Концептуально это находит отражение в реализации дифференцированного подхода к кибербезопасности и назначению уровней кибербезопасности для активов АСУТП АЭС и интеграции уровней кибербезопасности в модель разграничения доступа, определяемой политикой безопасности системы.

5.2. Дифференцированный подход и разграничение доступа

Реализация архитектуры кибербезопасности основывается на дифференцированном подходе, при котором критическим активам объекта назначаются уровни в соответствии с их критичностью для безопасности АЭС. Между классифицированными активами, согласно принципу глубоко эшелонированной защиты, реализуются соответствующие барьеры, препятствующие распространению угрозы. Каждому активу назначается набор мер для предупреждения, выявления и реагирования на кибератаки. Степень защиты каждого из активов должна соответствовать присвоенному для него уровню кибербезопасности.

Одна из проблем, которая решается при проектировании архитектуры заключается в установлении соответствия уровней кибербезопасности с иерархией (системой уровней) доступа. Уровень доступа, присваиваемый активу, однозначно определяется структурой информационных связей, существующих в системе. Правильно спроектированная система должна иметь однозначное и прямое соответствие между уровнями кибербезопасности и доступа, что обеспечит отсутствие слабых мест в защите системы и, как следствие, снижение риска кибербезопасности.

Существует два основных подхода к заданию уровней доступа — это рассмотрение доминирования и информационных потоков. Для основных целей проектирования архитектуры кибербезопасности оба подхода фактически одинаковы и различаются лишь семантикой проблемы. Далее рассмотрим более понятный для специалистов АСУТП подход информационных потоков.

В соответствии с этим подходом:

- каждому активу присваивается уровень доступа (или метка безопасности), и передача информации между активами подразумевает передачу информации между соответствующими уровнями доступа;

— определен оператор объединения уровней $A \oplus B = C$, т. е. активы, содержащие информацию из уровней A и B , должны быть помечены уровнем C ;

— определено отношение $A \rightarrow B$, чтобы указать, что информация разрешена для потока от актива уровня доступа A до актива уровня доступа B .

Отношение $A \rightarrow B$ определяет отношение частичного упорядочения, посредством которого можно указать иерархию уровней доступа, и которое разбивает множество активов на непересекающиеся подмножества, упорядоченные так, что если актив входит в подмножество с номером i , то *следующий* за ним входит в подмножество с номером, большим чем i , такое отношение транзитивно и антисимметрично.

Определение 1. Полученные непересекающиеся подмножества называются уровнями доступа.

Определение 2. Актив x «выше», чем актив y , если $A \rightarrow B$ верно, а $B \rightarrow A$ ложно. В этом случае говорят, что B «ниже», чем A .

Определение 3. Два актива x и y находятся на одной высоте (или уровне), если x и y принадлежат одному уровню. Аналогично говорят, уровень доступа $L2$ выше, чем уровень $L1$, если для каждой вершины $l2 \in L2$ и для каждой вершины $l1 \in L1$, $l2 > l1$. Аналогично для случая, когда один уровень ниже другого. ♦

В основе известных архитектур кибербезопасности [11, 19] АСУТП АЭС, в том числе и рассмотренной далее, лежит реализация иерархии уровней доступа в виде решетки для свойств целостности (модель Биба [20]).

5.3. Архитектура кибербезопасности RG 5.71

Архитектура RG 5.71 [19] получила широкое распространение, и основные ее положения повторяются во многих документах, посвященных кибербезопасности. Данная архитектура, фактически с момента появления, стала стандартной для атомной отрасли в международном масштабе. Она с небольшими вариациями повторена в документах

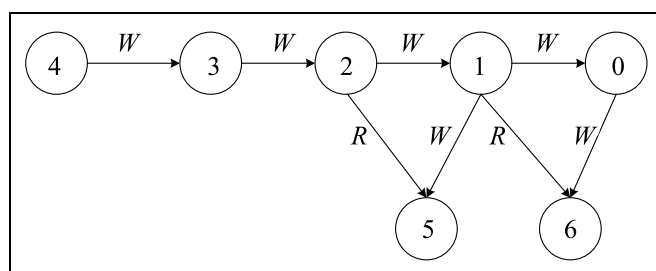


Рис. 4. Упрощенная информационная модель архитектуры безопасности RG 5.71: узлы графа 0—4 соответствуют субъектам; узлы 5 и 6 — объектам

МАГАТЭ и МЭК. Основные изменения относятся к числу уровней в системе и рекомендациям по менее строгому применению однонаправленных связей между активами на разных уровнях системы. Наличие различных вариантов обусловлено большой сложностью АСУТП АЭС как системы систем и сложившейся практикой работы.

Для архитектуры вводится 5 уровней компьютерной безопасности: уровни нумеруются с 4 по 0. Самый высокий (значимый) — 4, самый низкий — 0. Без ограничений поток данных разрешен с верхнего на более низкий уровень безопасности, передача данных с нижнего уровня на высший разрешена только для двух первых уровней кибербезопасности. Основная цель архитектуры кибербезопасности состоит в сохранении целостности данных и предотвращении модификации информации системами нижнего уровня кибербезопасности в системах верхнего уровня. Для достижения данной цели реализуется модель Биба для двух высших уровней кибербезопасности. Для низших уровней модель реализуется в ограниченном масштабе. Вводится запрет инициации соединения для нижнего уровня.

На рис. 4 приведена возможная информационная модель архитектуры типа RG 5.71. В модели для упрощения на каждом из уровней выделен единственный актив-субъект. Поток данных от верхнего уровня к нижнему определяется наличием доступа на запись (w) от субъекта, принадлежащего более высокому уровню, к низкому. Для безопасного доступа субъектов с уровнями 0 и 1 к субъектам на более высоких уровнях 1 и 2 соответственно, применяется однонаправленный доступ по чтению (r) через буферные активы-объекты 5 и 6 соответственно. Применение буферных активов необходимо для того чтобы запретить возможность инициации доступа с нижнего уровня на верхний.

Особенность архитектуры RG 5.71 заключается в наличии рекомендаций по формированию зональной модели, позволяющей группировать системы по важности для физической безопасности и безопасности объекта. При использовании зональной модели рекомендуется соблюдать рекомендации:

- каждая зона включает в себя системы, имеющие одинаковую или сопоставимую важность для физической безопасности и безопасности установки;
- системы, относящиеся к одной зоне, имеют аналогичные потребности в отношении защитных мер;
- различные компьютерные системы, принадлежащие одной зоне, образуют область надежной связи для внутренней коммуникации в пределах этой зоны;



— зональные границы требуют механизмов развязки для потоков данных на основе политики, зависящей от конкретной зоны;

— в целях улучшения конфигурации зоны могут быть разделены на субзоны.

В целом архитектура RG 5-71 представляет собой разумное сочетание сложившейся практики построения АСУТП АЭС с требованиями по кибербезопасности.

ЗАКЛЮЧЕНИЕ

Рассмотрены вопросы, связанные с процессом управления кибербезопасностью АСУТП АЭС. Анализируются практические вопросы оценки риска, методики назначения уровней кибербезопасности, меры защиты и построения архитектуры кибербезопасности. Обосновано утверждение, что основная цель киберзащиты АСУТП АЭС состоит в недопущении нарушения норм ядерной безопасности, и это определяет сильную связь классификации систем АСУТП АЭС по ядерной и кибербезопасности.

Сформулированы основные задачи обеспечения кибербезопасности АСУТП АЭС. Обозначены принципы построения архитектуры кибербезопасности. Рассмотрена связь между уровнями кибербезопасности и классическими уровнями доступа, известными из практики информационной безопасности. Приведен пример практической реализации пятиуровневой архитектуры кибербезопасности на основе модели Биба с разбиением уровней на отдельные зоны безопасности и предложена возможная информационная модель данной архитектуры.

ЛИТЕРАТУРА

1. Менгазетдинов Н.Э., Бывайков М.Е., Зуенков М.А. и др. Комплекс работ по созданию первой управляющей системы верхнего блочного уровня АСУТП для АЭС «Бушер» на основе отечественных информационных технологий. — М.: ИПУ РАН, 2013.
2. Федеральные нормы и правила в области использования атомной энергии «Требования к управляющим системам, важным для безопасности атомных станций» НП-026-16. Зарегистрировано в Минюсте России 14 декабря 2016 г., № 44712.
3. ГОСТ Р МЭК 61508-4. Функциональная безопасность систем электрических, электронных, программируемых электронных связанных с безопасностью. Часть 4. Термины и определения. 2012.
4. Менгазетдинов Н.Э., Полетыкин А.Г., Промыслов В.Г. Концепция обеспечения защиты от несанкционированного доступа АСУТП АЭС «Бушер-1» // Автоматизация в промышленности. — 2005. — № 5. — С. 3—5.
5. Промыслов В.Г., Полетыкин А.Г. Методы борьбы с киберугрозами АСУТП современного предприятия // Information Security / Информационная безопасность. — 2016. — № 1.

6. Харченко В., Скляр В., Брежнев Е. Безопасность информационно-управляющих систем и инфраструктур ISBN: 978-3-8473-9859-2 // Pa marium academic publishing, 2013.
7. Промыслов В.Г., Полетыкин А.Г. Формальная иерархическая модель безопасности верхнего уровня АСУТП АЭС // Ядерные измерительно-информационные технологии. — 2012. — Т. 4 (44). — С. 39—53.
8. IEC 62645 ed 1. Атомные электростанции. Системы контроля и управления. Требования к программам обеспечения безопасности для компьютерных систем. — 2014.
9. IEC 62859. Атомная электростанция. Приборы и системы управления. Требования к координации безопасности и кибербезопасности. — 2016.
10. Computer Security at Nuclear Facilities Technical Guidance Reference Manual IAEA Nuclear Security Series. — 2011. — N 17.
11. ФСТЭК России от 14.03.2014 № 31 (ред. от 23.03.2017) «Об утверждении требований к обеспечению защиты информации в автоматизированных системах управления производственными и технологическими процессами на критически важных объектах, потенциально опасных объектах, а также объектах, представляющую повышенную опасность для жизни и здоровья людей и для окружающей природной среды». Зарегистрировано в Минюсте России 30.06.2014, № 32919.
12. NST055 Handbook on the design of physical protection systems for nuclear material and nuclear facilities, (Draft) August, 2017.
13. РФ Федеральный закон от 27.07.2006 № 149-ФЗ (ред. от 19.12.2016) «Об информации, информационных технологиях и о защите информации».
14. ISO 13335-1 2004. Информационная технология. Методы обеспечения безопасности. Управление безопасностью информационных и телекоммуникационных технологий. Часть 1. Концепция и модели управления безопасностью информационных и телекоммуникационных технологий.
15. ISO/IEC 27005 Международный Стандарт. Информационная технология. Методы защиты. Менеджмент рисков информационной безопасности.
16. Постановление Госатомнадзора РФ от 14.11.1997 № 9 «Об утверждении и введении в действие нормативного документа ПНАЭ Г-01-011-97 «Общие положения обеспечения безопасности атомных станций. ОПБ-88/97».
17. IEC 6155134. Атомные электростанции. Системы контроля и управления, важные для безопасности. Общие требования.
18. IEC/TS 62443-1-1 Терминология, концепции и модели. — 2009.
19. US. Nuclear regulatory commission. Research regulatory guide 5.71 Cyber security programs for nuclear facilities, January 2010.
20. Девянин П.Н. Модели безопасности компьютерных систем. — М.: Academia, 2005. — 144 с.

Статья представлена к публикации членом редколлегии В.Г. Лебедевым.

Бабаев Денис Игоревич — нач. отдела, АО «НИИ по эксплуатации атомных электростанций», г. Москва, ✉ dibabaev@vniiaces.ru,

Полетыкин Алексей Григорьевич — д-р техн. наук, зав. лабораторией, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ poletik@ipu.ru,

Промыслов Виталий Георгиевич — канд. физ.-мат., вед. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ v1925@mail.ru,

Тимофеев Михаил Юрьевич — инженер-программист, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ timof-1964@mail.ru.

МЕТОД РЕКУРРЕНТНОГО И Z-РЕКУРРЕНТНОГО УПРАВЛЕНИЯ ФУНКЦИОНИРОВАНИЕМ СЛОЖНОЙ СИСТЕМЫ¹

А.Ф. Резчиков, В.А. Твердохлебов

Разработан метод построения правил управления взаимодействиями процессов функционирования сложной системы. Процессы рассмотрены как последовательности причинно-следственно связанных событий, заданных значениями показателей их свойств. Структура правил управления представлена на основе рекуррентных и впервые вводимых Z-рекуррентных определений последовательностей в форме функциональных зависимостей значений показателей свойств событий. Отмечено, что разработанные модели событий, процессов и функционирования сложной системы в целом применимы при решении задач управления, контроля и диагностирования.

Ключевые слова: функционирование сложной системы, процесс, событие, правило управления, рекуррентное и Z-рекуррентное определение последовательности.

ВВЕДЕНИЕ

В статье исследуются новые формы правил управления функционированием сложной системы с заданным целевым предназначением, в которой действуют и взаимодействуют разнородные процессы в соответствии с разработанной классификацией процессов, охватывающей 63 класса процессов $P_1, P_2, \dots, P_{63}$ (класс командно-информационных управляющих процессов P_1 , класс процессов действий человеческих звеньев P_2 , класс процессов действий неисправностей в технике и оборудовании P_3 и т. д. [1–3]). Математические модели функционирования системы и модели отдельных процессов определяются на основе функциональных связей событий в процессах, т. е. в последовательностях событий. Структуры последовательностей событий представляются с помощью рекуррентных и впервые вводимых Z-рекуррентных определений последовательностей. Классическое рекуррентное определение последовательности вида

$$\xi = \langle a_1, a_2, \dots, a_c \rangle, \quad c \in N^+, \quad (1)$$

посредством рекуррентной формы F^m порядка m

$$a_{t+m+1} = F^m(a_{t+1}, a_{t+2}, \dots, a_{t+m}),$$

где $0 \leq t \leq c - m - 1$, расширяется до Z-рекуррентного порядка $(i_1, i_2, \dots, i_{m_1}; j_1, j_2, \dots, j_{m_2})$ определения последовательности $\xi(1)$ с применением Z-рекуррентной формы вида $F^{Z,M}$.

$$\begin{aligned} &\langle a_{t+j_1}, a_{t+j_2}, \dots, a_{t+j_{m_2}} \rangle = \\ &= F^{Z,M}(a_{t+i_1}, a_{t+i_2}, \dots, a_{t+i_{m_1}}), \end{aligned} \quad (3)$$

где $M = (i_1, i_2, \dots, i_{m_1}; j_1, j_2, \dots, j_{m_2})$.

Рекуррентная F^m и Z-рекуррентная $F^{Z,M}$ формы [4, 5] переводят причинно-следственные связи событий в функциональные зависимости элементов в последовательности ξ и определяют структуры правил управления, т. е. правил, по которым в момент (интервал) времени t события в последовательности ξ , представленные формами F^m и $F^{Z,M}$ как совершившиеся события, определяют выбор управляющих воздействий с целью формирования событий a_{t+m+1} или $(a_{t+j_1}, a_{t+j_2}, \dots, a_{t+j_{m_2}})$. Ре-

¹ Работа выполнена при финансовой поддержке РФФИ, грант № 16-08-00697.



куррентная форма F^m порядка m определяет покрытие последовательности ξ подпоследовательностями $\xi_1, \xi_2, \dots, \xi_{c-m}$, каждая из которых имеет длину $m+1$, является сдвигом слева направо по отношению к непосредственно предшествующей подпоследовательности и представляет функциональную связь событий в моменты $t+1, t+2, \dots, t+m$ (причины) с событием в момент $t+m+1$ (следствием). Правила управления последовательностью событий ξ для каждой подпоследовательности ξ_v , $1 \leq v \leq c-m$, определяют конкретное управляющее воздействие на основе данных о m предшествующих событиях и данных о требуемом формировании $(m+1)$ -го события. Величина $(m+1)v$ характеризует объем данных в правиле управления для реализации правила по управлению всей последовательностью ξ . Введение Z -рекуррентного определения последовательности существенно и принципиально расширяет возможности для преобразования причинно-следственных связей событий в функциональные зависимости. На числовые показатели $i_1, i_2, \dots, i_{m_1}, j_1, j_2, \dots, j_{m_2}$ в порядке Z -рекуррентной формы для последовательности ξ (1) налагаются ограничения:

$$i_1 < i_2 < \dots < i_{m_1}, \quad j_1 < j_2 < \dots < j_{m_2};$$

$$1 \in \{i_1, i_2, \dots, i_{m_1}, j_1, j_2, \dots, j_{m_2}\};$$

$$\tau_{\xi}^1 \leq c, \quad \tau_{\xi}^1 = \min\{i_1, i_2, \dots, i_{m_1}, j_1, j_2, \dots, j_{m_2}\}.$$

Покрытие последовательностью ξ , соответствующее рекуррентному определению последовательности, является простейшим. Предлагаемое Z -рекуррентное определение последовательности принципиально расширяет возможности для определения структуры правила управления, так как применение Z -рекуррентного покрытия может не только уменьшать размеры правила управления, но и (возможно) исключать из анализа данные о трудно определяемых событиях. В правилах управления функционированием реальной сложной системой должны быть учтены и представлены:

- целевое предназначение системы;
- общая схема этапов рассматриваемой реализации целевого предназначения системы;
- варианты функционирования системы в работоспособном состоянии;
- множество учитываемых процессов, в которых определены изменения событий, связанных с ошибками в командно-информационных процессах, с ошибками в процессах действий человеческих звеньев, с неисправностями в технике и обо-

рудовании и др., для процессов из классов процессов $P_1, P_2, \dots, P_{63}$;

— в моделях для каждого процесса значения показателей свойств $R_1, R_2, \dots, R_k$ событий, характеризующих работоспособное функционирование системы и функционирование системы при наличии ошибок, неисправностей, дефектов и др. в процессах.

Рекуррентное и Z -рекуррентное определения процессов как последовательностей событий в функционировании системы позволяют компактно и систематизированно представлять функциональные зависимости между значениями показателей свойств событий-причин и событий-следствий. Для практического применения полученных результатов по рекуррентному и Z -рекуррентному управлению исходной информацией служит множество $H = H_1 \cup H_2$ вариантов функционирования системы по целевому предназначению, включающее в себя учитываемые варианты работоспособного функционирования (множество H_1) и функционирования с ошибками, неисправностями, дефектами (множество H_2):

$$H_1 = \{ \langle a_{11}, a_{12}, \dots, a_{1c_1} \rangle, \langle a_{21}, a_{22}, \dots, a_{2c_2} \rangle, \dots, \langle a_{\alpha 1}, a_{\alpha 2}, \dots, a_{\alpha c_{\alpha}} \rangle \}, \quad (2)$$

$$H_2 = \{ \langle a_{11}, a_{12}, \dots, a_{1d_1} \rangle, \langle a_{21}, a_{22}, \dots, a_{2d_2} \rangle, \dots, \langle a_{\beta 1}, a_{\beta 2}, \dots, a_{\beta d_{\beta}} \rangle \}. \quad (3)$$

Для реальных систем множества вариантов работоспособного функционирования и функционирования с ошибками, неисправностями, дефектами неограниченные и, следовательно, применение предлагаемого метода управления функционированием сложной системы позволяет получить результаты, ограниченные выбором множеств H_1 и H_2 . Для обеспечения эффективности предлагаемого метода рекуррентного и Z -рекуррентного управлений функционированием сложной системы множества H_1 и H_2 должны разбиваться на подмножества в соответствии с интерпретацией последовательностей событий на основе классификации процессов на классы $P_1, P_2, \dots, P_{63}$. При классификации функционирования системы естественно выделять:

- класс вариантов функционирования с полным выполнением целевого предназначения;
- класс вариантов функционирования с частичным выполнением целевого предназначения;
- класс вариантов функционирования без частичного и полного выполнения целевого предназначения.

Конкретный вариант функционирования системы рассматривается как перевод системы из заданного начального состояния в заданное заключительное состояние системы с наблюдением событий, составляющих процессы в системе, и управлением событиями и процессами. Элементы $a_{i,j}$ и $b_{v,\mu}$ в последовательностях, представленных в множествах H_1 (2) и H_2 (3), являются математическими моделями событий. В приложениях следует различать теоретические формы моделей событий (обозначения буквами, представление конкретными значениями конкретных свойств) и реально, фактически наблюдаемую средствами контроля и диагностирования модель события в форме наборов значений показателей свойств. В данной статье, в основном, используются теоретические модели событий, так как фактически наблюдаемые средствами контроля и диагностирования значения показателей свойств потребуют рассмотрения вариантов возможностей средств контроля и диагностирования. Целевое предназначение системы включает в себя перевод из начального состояния в заключительное с выполнением специфических задач. Например, в функционировании системы управления полетом воздушного судна (ВС) можно обобщенно и условно выделить этапы [6]:

- движение ВС по рулежной дорожке до торца взлетно-посадочной полосы (ВПП),
- разбег ВС по ВПП до достижения скорости отрыва ВС от ВПП,
- движение ВС до набора высоты эшелона полета,
- полет ВС по маршруту с возможным выполнением специфических задач,
- полет ВС до выхода на посадочную глиссаду,
- снижение по посадочной глиссаде до касания ВПП и торможение,
- движение ВС по рулежной дорожке до стоянки.

В общем правиле управления движением ВС для каждого из перечисленных этапов применяются частные правила, в которых представлена специфика этапа. Если управление движением ВС представлять как управление событиями в процессах и процессах, то события в последовательностях событий из множества $H = H_1 \cup H_2$ находятся в причинно-следственных связях, в которых события-причины могут включать в себя: ошибки в классах процессов P_1 и P_2 ; неисправности в технике и оборудовании (класс процессов P_3); нарушения в энергообеспечении (класс процессов P_4); проявившиеся недостатки в обеспечении функционирования системы комплектующими, сырьем, грузами и т. д. (класс процессов P_5); воздействие

внешней среды, включая действие метеоусловий. В последовательностях событий из множества H включаются процессы, представленные классами процессов $P_7, P_8, \dots, P_{63}$, свойства которых определяются взаимодействиями шести базовых процессов $P_1, P_2, \dots, P_6$ в сочетаниях по два, три, ..., всех шести.

В статье содержатся результаты исследования структур правил управления системами. Структуры правил определяются на основе применения форм рекуррентного и Z -рекуррентного определений последовательности событий, образующих в соответствии с причинно-следственными связями событий процессы в системе. Исследование опирается на ранее разработанные авторами классификацию процессов, содержащую 63 класса процессов и алгебру причинно-следственных комплексов [1–3, 7]). К новым результатам относятся метод преобразования структур причинно-следственных связей событий и функциональные структуры, разработанные на основе рекуррентных и Z -рекуррентных последовательностей событий.

1. ОПРЕДЕЛЕНИЕ СОБЫТИЙ ЗНАЧЕНИЯМИ ПОКАЗАТЕЛЕЙ ИХ СВОЙСТВ

Пусть последовательность ξ (1) интерпретируется как последовательность событий, составляющая процесс функционирования системы на некотором этапе ее функционирования. Для преобразования причинно-следственных связей событий в функциональные зависимости значений показателей свойств событий по набору свойств $R_1, R_2, \dots, R_k$ событий требуется определить области событий-причин для каждого события-следствия. Это можно сделать на основе рекуррентного определения последовательности ξ рекуррентной формой F^m порядка m , которая определяет функциональную зависимость f :

$$\begin{aligned} f(L(a_1(0)), L(a_2(0)), \dots, L(a_m(0))) &= L(a_{m+1}(0)), \\ f(L(a_2(1)), L(a_3(1)), \dots, L(a_{m+1}(1))) &= L(a_{m+2}(1)), \\ f(L(a_{c-m}(c-m-1)), L(a_{c-m+1}(c-m-1)), \dots, \\ &L(a_{c-1}(c-m-1))) = L(a_c(c-m-1)), \end{aligned} \quad (4)$$

где $L(a_j(t))$ — набор значений показателей свойств события a_j в момент времени t . Формулы (4) определяют правила выбора и вычисления управляющих воздействий для формирования события $a_j(t+m+1)$ по показателям m непосредственно предшествующих событий.

Принципиально большие возможности представляет Z -рекуррентное определение последо-

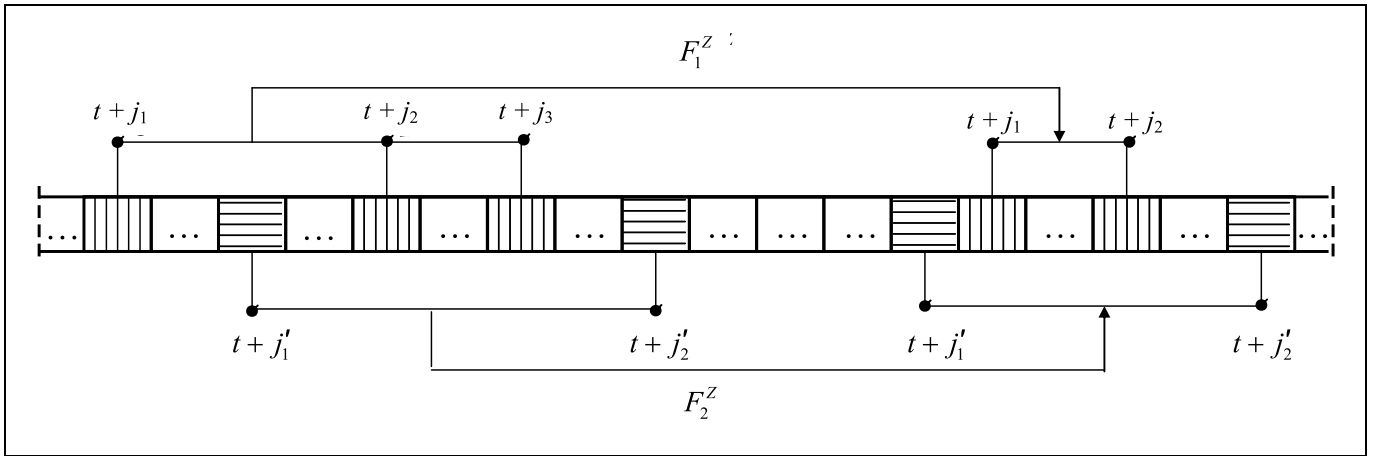


Рис. 1. Пример Z-рекуррентных форм F_1^Z и F_2^Z порядков $(\{i_1, i_2, i_3\}; \{j_1, j_2\})$ и $(\{i'_1, i'_2\}; \{j'_1, j'_2\})$ для декомпозиции последовательности на две Z-рекуррентно определенных подпоследовательности (для случая t -го момента времени)

вательности ξ с помощью Z-рекуррентного определения последовательности ξ (1) с применением Z-рекуррентной формы вида $F^{Z,M}$ порядка $M = (i_1, i_2, \dots, i_v; j_1, j_2, \dots, j_\mu)$. В этом случае причинно-следственные связи представляются функциональной зависимостью f в момент времени t событий-следствий $a_{t+j_1}, a_{t+j_2}, \dots, a_{t+j_\mu}$ от событий-причин $a_{t+i_1}, a_{t+i_2}, \dots, a_{t+i_v}$. Получаем функциональные зависимости f_Z значений показателей свойств событий-следствий от значений показателей свойств событий-причин:

$$\begin{aligned} f_Z(L(a_{i_1}(0)), L(a_{i_2}(0)), \dots, L(a_{i_v}(0))) &= \\ &= (L(a_{j_1}(0)), L(a_{j_2}(0)), \dots, L(a_{j_\mu}(0))), \\ f_Z(L(a_{i_1+1}(1)), L(a_{i_2+1}(1)), \dots, L(a_{i_v+1}(1))) &= \\ &= (L(a_{j_1+1}(1)), L(a_{j_2+1}(1)), \dots, L(a_{j_\mu+1}(1))), \\ f_Z(L(a_{i_1+t_0}(t_0-1)), L(a_{i_2+t_0}(t_0-1)), \dots, \\ L(a_{i_v+t_0}(t_0+1))) &= (L(a_{j_1+t_0}(t_0-1)), \\ L(a_{j_2+t_0}(t_0-1)), \dots, L(a_{j_\mu+t_0}(t_0-1))), \end{aligned} \quad (5)$$

где $t_0 = c - \tau_\xi^2$, $\tau_\xi^2 = \max\{i_1, i_2, \dots, i_{m_1}, j_1, j_2, \dots, j_{m_2}\}$.

С помощью нескольких Z-рекуррентных форм можно проводить декомпозицию последовательности событий на набор подпоследовательностей, извлекаемых из последовательности с возможными пропусками в порядке расположения в порядке последовательности. При такой декомпозиции можно представлять причинно-следственные связи событий с разделением на события, подготавлива-

ющие при совершении предшествующего события условия для формирования отдаленных событий. Строгое определение таких связей определяется порядком $(\{i_1, i_2, \dots, i_v\}; \{j_1, j_2, \dots, j_\mu\})$. На рис. 1 показан пример декомпозиции последовательности событий на две подпоследовательности, определяемые Z-рекуррентными формами F_1^Z и F_2^Z порядков $(\{i_1, i_2, i_3\}; \{j_1, j_2\})$ и $(\{i'_1, i'_2\}; \{j'_1, j'_2\})$ соответственно.

Наборы (4) и (5) фактически представляют собой структуры дискретных детерминированных правил управления для ориентации — какие значения показателей свойств событий следует сформировать при управлении и какие значения показателей свойств событий имеются как данные и исходные. Последовательности событий, определяющие в наборах (4) и (5) процессы функционирования системы, могут быть систематизированы в общую структуру в форме сети, которую для наглядности и удобства применения будем представлять в виде дерева. В каждом событии в момент времени t , входящим в последовательность событий, составляющих функционирование системы, представлены события, относящиеся к конкретным процессам из классов процессов $P_1, P_2, \dots, P_{63}$. Структура такого события a_t , $1 \leq t \leq c$, в последовательности $\xi = \langle a_1, a_2, \dots, a_c \rangle$, $c \in N^+$, может быть представлена как совмещение трех компонентов:

a_t^1 — событие, определяющее действия внешней среды в событии a_t в момент времени t ;

a_t^2 — событие, определенное предшествующими событиями $a_{t-m}, a_{t-m+1}, \dots, a_{t-1}$;

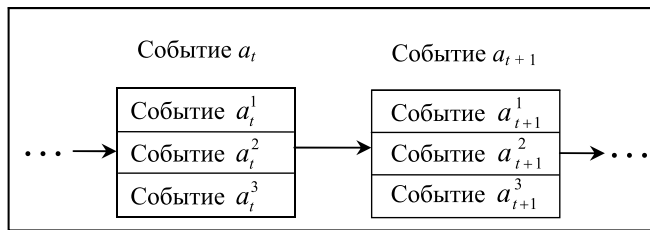


Рис. 2. Структура причинно-следственной связи событий a_t и a_{t+1} с включением действий внешней среды, предшествующих событиям и событиям, представляющих изменения процессов в системе

a_t^3 — событие, представляющее действия в событии a_t ошибок, неисправностей, нарушений и других факторов в процессах в момент времени t .

Представление события a_t в форме $a_t = \{a_t^1\} \cup \{a_t^2\} \cup \{a_t^3\}$ позволяет рассматривать событие a_t как следствие причинно-следственной связи с действием внешней среды и действием причин, возникающих в связи с изменениями процессов в системе (рис. 2).

Управление функционированием сложной системы, включающее в себя решение задач контроля и диагностирования, а также решение задач восстановления работоспособного функционирования разработано в форме общей структуры причинно-следственных связей событий в процессах и представлением компонентов, соответствующих рекуррентным и Z-рекуррентным определениям последовательностей событий. На основе рекуррентных и Z-рекуррентных определений последовательностей в общей структуре причинно-следственных связей событий определяются частные правила управления подпоследовательностями событий и переходы от одних частных правил к другим частным правилам в соответствии с возникновением реальных событий в процессе функционирования сложной системы.

2. ДИСКРЕТНЫЕ ДЕТЕРМИНИРОВАННЫЕ МОДЕЛИ СОБЫТИЯ, ПРОЦЕССА И ФУНКЦИОНИРОВАНИЯ СЛОЖНОЙ СИСТЕМЫ

Формальный аппарат рекуррентного и Z-рекуррентного определений последовательностей позволяет представлять модели события, процесса и функционирования сложной системы на основе функциональных зависимостей между элементами последовательностей. В модель функционирования системы будем включать статические модели событий как срезы в моменты времени процессов и динамические модели процессов как последовательности событий. Построение моделей для ре-

альных сложных систем ограничивается выбором для рассмотрения и включения в модель:

- набора процессов $\{P_{1,1}, P_{1,2}, \dots, P_{1,n_1}\}, \{P_{2,1}, P_{2,2}, \dots, P_{2,n_2}\}, \dots, \{P_{63,1}, P_{63,2}, \dots, P_{63,n_{63}}\}$ из классов процессов $P_1, P_2, \dots, P_{63}$;
- набора свойств $R_1, R_2, \dots, R_k$ представляемых в моделях процессов;
- набора множеств $W_1, W_2, \dots, W_k$ значений показателей свойств $R_1, R_2, \dots, R_k$;
- множества $E = E_1 \cup E_2$ вариантов функционирования реальной сложной системы в работоспособном состоянии (множество вариантов E_1) и при дефектах (множество вариантов E_2).

Большой объем этих исходных данных для построения моделей является следствием сложности системы, выбранных полноты и точности моделей, а также предполагаемых к реализации условий управления функционированием системы. Основной формой модели события служит таблица, в которой представлены значения показателей свойств для конкретных процессов, расположенных в строках, и для конкретных свойств, расположенных в столбцах. В табл. 1 каждый класс процессов $P_i, 1 \leq i \leq 63$, представлен срезом в момент t конкретных процессов $\{P_{i,1}, P_{i,2}, \dots, P_{i,n_i}\}$, относящихся к объектам или субъектам, участвующим в процессе из класса P_i .

Таблица 1

Табличная модель процесса функционирования в конкретный момент времени t

$T(t)$		R_1	...	R_k
$P_1(t)$	$P_{1,1}(t)$	$w_{1,1,1}(t)$	...	$w_{1,1,k}(t)$
	$P_{1,2}(t)$	$w_{1,2,1}(t)$	...	$w_{1,2,k}(t)$
	...	...	...	...
	$P_{1,n_1}(t)$	$w_{1,n_1,1}(t)$	...	$w_{1,n_1,k}(t)$
$P_2(t)$	$P_{2,1}(t)$	$w_{2,1,1}(t)$	...	$w_{2,1,k}(t)$
	$P_{2,2}(t)$	$w_{2,2,1}(t)$	...	$w_{2,2,k}(t)$
	...	...	...	...
	$P_{2,n_2}(t)$	$w_{2,n_2,1}(t)$	...	$w_{2,n_2,k}(t)$
...	...	...	...	...
$P_{63}(t)$	$P_{63,1}(t)$	$w_{63,1,1}(t)$	...	$w_{63,1,k}(t)$
	$P_{63,2}(t)$	$w_{63,2,1}(t)$	...	$w_{63,2,k}(t)$
	...	...	...	...
	$P_{63,n_{63}}(t)$	$w_{63,n_{63},1}(t)$	...	$w_{63,n_{63},k}(t)$



Например, процессами из класса P_2 действий человеческих звеньев в управлении полетом ВС будут процессы действий командира ВС ($P_{2,1}$), действий второго пилота ($P_{2,2}$), действий диспетчера ($P_{2,3}$) и т. д. В табл. 1 процесс представляется значениями показателей свойств события в момент времени t . Моделью процесса полагаются функциональные связи значений показателей свойств, представленных в таблицах вида табл. 1 для моментов времени t и $t + 1$. Такую связь можно формально представить таблицами вида табл. 2, в которых для каждого процесса P_{ij} явно определяются связи значений показателей свойств соответствующих событий в моменты времени t и $t + 1$. В табл. 2 для конкретного процесса P_{ij} и множеств $W_1, W_2, \dots, W_k$ значений показателей свойств $R_1, R_2, \dots, R_k$ определены изменения значений свойств событий по правилу: величина $w_{ij\nu\mu}(t)$, представляющая μ -е значение (в множестве $W_{ij\nu}$) свойства R_ν в процессе P_{ij} заменяется величиной $h_{ij\nu}(w_{ij\nu\mu}(t))$. Функции $h_{ij1}, h_{ij2}, \dots, h_{ijk}$, как показывает классификация процессов на классы $P_1, P_2, \dots, P_{63}$, должны быть различны по математической природе: в форме дифференциальных и интегральных уравнений, логических уравнений, таблиц, матриц, графов и др.

Символьные обозначения событий в множествах H_1 и H_2 заменим их табличными представлениями в форме таблиц вида табл. 1 и введем соответствующие полученные последовательности:

$$H'_1 = \{ \langle a'_{11}, a'_{12}, \dots, a'_{1c_1} \rangle, \langle a'_{21}, a'_{22}, \dots, a'_{2c_2} \rangle, \dots, \langle a'_{\alpha 1}, a'_{\alpha 2}, \dots, a'_{\alpha c_\alpha} \rangle \}, \quad (6)$$

$$H'_2 = \{ \langle b'_{11}, b'_{12}, \dots, b'_{1d_1} \rangle, \langle b'_{21}, b'_{22}, \dots, b'_{2d_2} \rangle, \dots, \langle b'_{\beta 1}, b'_{\beta 2}, \dots, b'_{\beta d_\beta} \rangle \}. \quad (7)$$

Табл. 2 как составляющая часть модели процесса P_{ij} используется по следующему правилу: если событие в момент t задается набором значений $(w_{ij1r}(t), w_{ij2u}(t), \dots, w_{ijkq}(t))$, где $r \in \{1, 2, \dots, m_1\}$, $u \in \{1, 2, \dots, m_2\}$, ..., $q \in \{1, 2, \dots, m_k\}$, то событие в момент $t + 1$ определяется набором $(h_{ij1}(w_{ij1r}(t)), h_{ij2}(w_{ij2u}(t)), \dots, h_{ijk}(w_{ijkq}(t)))$.

Исследования показали, что для управления функционированием сложной системы, в котором исходными данными являются целевое предназначение системы и решаемая задача, варианты функционирования системы в работоспособном состоянии и функционирования системы при выбранном множестве возможных дефектов в процессах, происходящих в системе, требуется последовательное построение большого числа моделей: моделей

Таблица 2

Табличная модель функциональной связи значений показателей свойств для событий в моменты времени t и $t + 1$

			R_1	R_2	...	R_k
$P_{ij}(t)$	R_1	$w_{ij11}(t)$ $w_{ij12}(t)$... $w_{ij1m_1}(t)$	$h_{ij1}(w_{ij11}(t))$ $h_{ij1}(w_{ij12}(t))$... $h_{ij1}(w_{ij1m_1}(t))$			
	R_2	$w_{ij21}(t)$ $w_{ij22}(t)$... $w_{ij2m_2}(t)$		$h_{ij2}(w_{ij21}(t))$ $h_{ij2}(w_{ij22}(t))$... $h_{ij2}(w_{ij2m_2}(t))$		
	...	...	...	...	...	...
	R_k	$w_{ijk1}(t)$ $w_{ijk2}(t)$... $w_{ijkm_k}(t)$				$h_{ijk}(w_{ijk1}(t))$ $h_{ijk}(w_{ijk2}(t))$... $h_{ijk}(w_{ijkm_k}(t))$

событий, моделей процессов, моделей функционирования системы. Из этих моделей конструируется фактическое управление функционированием системы, в котором последовательность событий доопределяется событиями, действующими извне (состояние метеоусловий, действия других систем и др.), и событиями, изменяющими процессы в системе (ошибки, неисправности, нарушения и другие факторы в процессах из классов процессов $P_1, P_2, \dots, P_{63}$).

3. МОДЕЛЬ СЛОЖНОЙ СИСТЕМЫ ДЛЯ УПРАВЛЕНИЯ ФУНКЦИОНИРОВАНИЕМ ПО ЦЕЛЕВОМУ ПРЕДНАЗНАЧЕНИЮ

Предложенные модели событий и процессов позволяют построить модель функционирования системы на основе совмещения событий и процессов. Если представлять модель функционирования системы графом с вершинами, интерпретируемыми как события, и дугами, рассматриваемыми как определения причинно-следственных связей событий, то получаем структуру, которую для наглядности будем представлять графом в виде дерева. Предполагается, что целевое предназначение системы и решаемые задачи определены соответствующими процессами, взаимосвязями и взаимодействиями процессов в форме множеств последовательностей событий H_1 и H_2 (см. выражения (2), (3)). События, представленные в последовательностях, определены моделями в форме таблиц вида табл. 1, а процессы в форме совмещений в форме таблиц вида табл. 2. Построение моделей событий и процессов требует решения задач разработки моделей учитываемых процессов из классов процессов $P_1, P_2, \dots, P_{63}$. Такой большой объем данных и моделей для построения модели управления функциониро-

ванием системы — непреложное следствие сложности задачи управления функционированием системы. В настоящей статье представлены результаты теоретических и общеметодических исследований и для разработки управления функционированием конкретной сложной системой по целевому предназначению дана общая содержательная интерпретация моделей событий, процессов и функционирования системы в целом.

Пусть информацией для управления функционированием системы являются множества, полученные из множеств H'_1 (6) и H'_2 (7) с учетом интерпретации элементов последовательностей событий в соответствии со значениями показателей свойств (4) и (5).

Для последовательностей из множеств H'_1 и H'_2 построим рекуррентные и Z-рекуррентные определения. Рассмотрим случай последовательности $\gamma_j = \langle a'_{j1}, a'_{j2}, \dots, a'_{jc_j} \rangle$ из множества H'_1 . Пусть для последовательности γ_j рекуррентная форма F_j^m имеет порядок m и определяет бинарное отношение

$$\begin{aligned} &(\langle a'_{j1}, a'_{j2}, \dots, a'_{jm} \rangle, a'_{jm+1}) \\ &(\langle a'_{j2}, a'_{j3}, \dots, a'_{jm+1} \rangle, a'_{jm+2}) \\ &\dots \\ &(\langle a'_{jc_j-m}, a'_{jc_j-m+1}, \dots, a'_{jc_j-1}, a'_{jc_j} \rangle). \end{aligned} \quad (8)$$

Рекуррентная форма и бинарное отношение (4) для последовательности событий γ_j определяет правило G_j управления событиями в последовательности γ_j . С использованием рекуррентных определений последовательностей множеству H'_1 последовательностей сопоставляется последовательность правил $G_a = \langle G_{a1}, G_{a2}, \dots, G_{ac_a} \rangle$. По построению правило G_{aj} определяет выбор управляющих воздействий в последовательности событий γ_j . Последовательности вида γ_j , $1 \leq j \leq c_a$, пересекаются по событиям, что усложняет общее правило G управления событиями в функционировании СЧМС.

Для управления последовательностью событий $\xi = \langle a_1, a_2, \dots, a_c \rangle$, $c \in N^+$, с использованием рекуррентного определения последовательности ξ , определяющая рекуррентная форма F^m порядка m задает покрытие, которое представлено бинарным отношением (8). На рис. 3 приведена схема связи правил управления $G_{a,\omega}$ и $G_{a,\omega}^*$, которые представляют работоспособное функционирование систе-

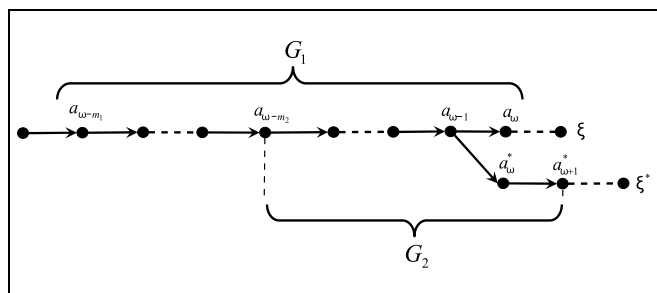


Рис. 3. Схема совмещения правила G_1 для управления последовательностью событий ξ с правилом G_2 для управления последовательностью событий ξ^* при определении последовательности ξ рекуррентной формой порядка m_1 и при определении последовательности ξ^* рекуррентной формой порядка m_2



мы и функционирование системы с дефектом в событии в момент $t = \omega$. Событиям a_ω и a_ω^* предшествует одна и та же последовательность событий, которая продолжается случаем $a_\omega \neq a_\omega^*$, в связи с чем для последовательностей событий G_1 и G_2 должно выполняться условие $G_1 \neq G_2$. Порядки m_1 и m_2 рекуррентных форм $F_1^{m_1}$ и $F_2^{m_2}$, рекуррентно определяющих последовательности ξ и ξ^* , не совпадают.

Метод рекуррентного и Z-рекуррентного управлений функционированием системы в работоспособном состоянии и при дефектах процессов в системе предполагает выбор:

- множества D_0 вариантов работоспособного функционирования системы и его определение;

- каждого класса процессов P_i , $1 \leq i \leq 63$, набора процессов $P_{i,1}, P_{i,2}, \dots, P_{i,n_i}$, представляющих класс P_i в функционировании системы;

- для множеств $M_1, M_2, \dots, M_{63}$ дефектов, процессов $P_{i,1}, P_{i,2}, \dots, P_{i,n_i}$, $1 \leq i \leq 63$, и их определение, представляющих выбранные процессы из классов $P_1, P_2, \dots, P_{63}$;

- свойств $R_1, R_2, \dots, R_k$ и соответствующих им множеств $W_1, W_2, \dots, W_k$ значений показателей свойств, которыми в форме табл. 1 будут определяться события;

- математических средств (дифференциальных и интегральных уравнений, логических уравнений, таблиц, матриц, графов) и их разработка для определения зависимостей значений показателей свойств в звеньях причинно-следственно связанных событий в процессах. Такие зависимости представляются таблицами вида табл. 2;

- рекуррентного или Z-рекуррентного определений последовательностей событий в процессах и в вариантах функционирования системы для построения правил управления и совмещения частных правил в общее правило управления функционированием системы.

Осуществление такого выбора требует исследований специфики рассматриваемой сложной системы. Общее правило управления определяется как композиция частных правил, определяемых на основе выбранного рекуррентного или Z-рекуррентного определений последовательностей событий в процессах и в вариантах функционирования системы (рис. 3). Структура общего правила управления G представима в форме дерева, в котором каждый путь $\eta = \langle G_{\eta_1}, G_{\eta_2}, \dots, G_{\eta_{u_\eta}} \rangle$ из корня в висящую вершину определяет последователь-

ность частных правил управления. Правило $G_{\eta_{j_\eta}}$, $1 \leq j_\eta \leq u_\eta$, соответствует части пути η , выделенной при рекуррентном (или Z-рекуррентном) определении последовательности η .

Выбор в момент времени t управляющего воздействия для формирования события a_t согласно общему правилу управления функционированием системы, представленного в форме дерева, осуществляется на основе данных о событиях $a_{t-m}, a_{t+m+1}, \dots, a_{t+1}$ при рекуррентном управлении. В случае, когда правила управления функционированием системы определяются с применением Z-рекуррентного определения последовательностей, рассматриваются функциональные зависимости между значениями показателей свойств: для значений показателей свойств событий $a_{t+i_1}, a_{t+i_2}, \dots, a_{t+i_{m_1}}$ определяются функциональные зависимости с значениями показателей свойств событий $a_{t+j_1}, a_{t+j_2}, \dots, a_{t+j_{m_2}}$.

ЗАКЛЮЧЕНИЕ

Рекуррентные модели и методы находят широкое применение в различных областях приложений (например, при оценке параметров движения сложных объектов [8], в задачах анализа систем с конечным числом частиц [9], в задачах управления рисками в инновационной деятельности [10], в навигации космических аппаратов [11], в теории стохастической фильтрации [12] и др.). В настоящей работе управление функционированием сложной системы рассматривается в традиционной форме перевода в пространстве состояний системы из заданных начального состояния в заключительное состояние.

Функционирование системы рассматривается как последовательность событий, формируемых с помощью управляющих воздействий, при которой учитываются действия внешней среды и действия дефектов процессов в системе. Данный подход опирается на ранее разработанные классификацию процессов [7, 13, 14], представление процессов последовательностями событий и представление событий значениями показателей свойств. Причинно-следственные связи событий в процессах преобразуются в функциональные зависимости между значениями показателей свойств событий-причин и событий-следствий. Общее правило управления функционированием системы для каждого варианта функционирования системы определяет последовательность частных правил, в которой каждое частное правило соответствует управлению для части последовательности причинно-следст-

венно связанных событий, вычисленной на основе рекуррентного или Z -рекуррентного определения последовательности событий. Применение рекуррентного и Z -рекуррентного определений последовательностей событий разработано и предлагается впервые. Рекуррентные и Z -рекуррентные формы в рекуррентных и Z -рекуррентных определениях последовательностей выделяют области значений показателей свойств событий с ограничением областей требующихся значений. Этим устанавливаются размеры правил управления процессами и функционированием системы в целом.

Сформулируем полученные результаты.

- На основе ранее разработанной классификации процессов на классы $P_1, P_2, \dots, P_{63}$ [1–3] и исследований структур рекуррентных и впервые вводимых Z -рекуррентных определений последовательностей событий [4] построены дискретные детерминированные модели событий (в форме таблиц связей событий со значениями показателей свойств $R_1, R_2, \dots, R_k$), модели процессов в виде последовательностей событий и модели функционирования сложных систем в форме структур, определяющих причинно-следственные связи событий и причинно-следственные связи процессов.
- Показано, что разработанная модель структуры причинно-следственных связей событий и причинно-следственных связей процессов определяет структуру общего правила управления функционированием сложной системы на основе частных правил управления конкретными процессами.
- Показано, что на уровне предлагаемых формальных моделей событий, процессов и функционирования сложной системы представлены действия внешней среды, предшествующие событиям и изменения процессов в функционировании системы, вызванные ошибками, неисправностями и другими факторами процессов. Это означает, что модель функционирования системы может использоваться для управления системой с включением в управление решений задач контроля, диагностирования и восстановления работоспособности.
- Впервые для решения задач контроля и диагностирования процессов функционирования сложных систем предлагается использовать равенства и неравенства числовых показателей рекуррентных и Z -рекуррентных определений последовательностей событий. В таких решениях задач работоспособные и неработоспособные процессы распознаются на основе изменений их причинно-следственных зависимостей от предшествующих событий.

Полученные результаты имеют теоретическое и общеметодическое значение, которые применимы в приложениях на основе достаточно точной и пол-

ной содержательной интерпретации моделей событий, процессов и функционирования системы.

ЛИТЕРАТУРА

1. Резчиков А.Ф., Твердохлебов В.А. Принцип причинно-следственной декомпозиции динамических систем: Монография. — Саратов: Изд. центр «Наука», 2013. — 56 с.
2. Резчиков А.Ф., Твердохлебов В.А. Причинно-следственные модели производственных систем. — Саратов: Научная книга, 2008. — 137 с.
3. Резчиков А.Ф., Твердохлебов В.А. Причинно-следственные комплексы взаимодействий в производственных процессах // Проблемы управления. — 2010. — № 3. — С. 51–60.
4. Твердохлебов В.А. Геометрическая форма автоматных отображений, рекуррентное и Z -рекуррентное определение последовательностей // Изв. Саратовского ун-та. Сер. Математика. Механика. Информатика. — 2016. — Т. 16, вып. 2. — С. 232–241.
5. Твердохлебов В.А. Z -рекуррентное определение последовательностей в задачах контроля и диагностирования процессов в системах // Доклады академии военных наук. — 2016. — № 2 (70). — С. 43–47.
6. Новожилов Г.В., Неймарк М.С., Цесарский Л.Г. Безопасность полета самолета. Концепция и технология. — М.: Машиностроение, 2003. — 144 с.
7. Твердохлебов В.А. Управление потоками состояний крупномасштабных систем // Материалы восьмой междунар. конф. «Управление развитием крупномасштабных систем MLSD'2015» / ИПУ РАН. — М., 2015. — С. 333–335.
8. Алешкин А.П., Архипова И.Г., Полиенко В.Н. и др. Метод рекуррентного оценивания параметров движения подводного объекта по данным космических навигационных определений буксируемой аппаратуры потребителя // Радиопромышленность. — 2018. — № 1. — С. 57–61.
9. Умирзаков И.Х. Рекуррентный метод определения различных наборов кластеров и распределения кластеров по размерам в системе с конечным числом частиц // Бутлеровские сообщения. — 2015. — Т. 44, № 10. — С. 45–63.
10. Брега Г.В. Рекуррентный подход к управлению рисками в инновационной деятельности // Управленческие науки. — 2015. — № 2. — С. 50–57.
11. Анцев Г.В., Лысенко Л.Н., Петров В.А. Повышение точности определения параметров орбит на основе применения операторов совмещения витковых оценок по результатам малоинтервальной обработки данных ГЛОНАСС // Вестник Моск. гос. техн. ун-та им. Н.Э. Баумана. Сер.: Приборостроение. — 2016. — № 5 (110). — С. 99–110.
12. Daum F. Nonlinear Filters: Beyond the Kalman Filter // IEEE Aerospace and Electronic Syst. — 2005. — Vol. 8. — P. 57–71.
13. Критические ситуации в человеко-машинных системах / под ред. А.Ф. Резчикова. — Саратов: Изд. центр «Наука», 2015. — 244 с.
14. Резчиков А.Ф., Твердохлебов В.А., Иващенко В.А. и др. Человек. Машина. Среда. — Саратов: Изд. центр «Наука», 2013. — 196 с.

Статья представлена к публикации членом редсовета чл.-корр. РАН П.П. Пархоменко.

Резчиков Александр Федорович — д-р техн. наук, чл.-корр. РАН, гл. науч. сотрудник, Институт проблем точной механики и управления РАН, г. Саратов; гл. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ✉ iptmuran@san.ru,

Твердохлебов Владимир Александрович — д-р техн. наук, гл. науч. сотрудник, Институт проблем точной механики и управления РАН, г. Саратов, ✉ tverdokhlebova@list.ru.

ГРАФОВАЯ МОДЕЛЬ РАСПРЕДЕЛЕНИЯ ЛОКОМОТИВОВ ДЛЯ ГРУЗОВЫХ ПЕРЕВОЗОК НА ЛИНЕЙНОМ УЧАСТКЕ ЖЕЛЕЗНОЙ ДОРОГИ. ЗАДАЧА О МАКСИМАЛЬНОМ ПО ВКЛЮЧЕНИЮ ПОКРЫТИИ ГРАФИКА¹

Л.Ю. Жиликова, Н.А. Кузнецов, В.Г. Матюхин, А.Б. Шабунин, А.К. Такмазьян

Работа посвящена формальной постановке и решению задачи, возникающей при распределении локомотивов между составами для осуществления грузовых железнодорожных перевозок с заранее заданным графиком. Цель состоит в том, чтобы определить, достаточно ли число локомотивов при заданном их начальном расположении для выполнения всех перевозок. Решение представлено в виде алгоритма, строящего покрытие графика: полное, если оно существует, или, в противном случае, частичное — максимальное по включению. Доказана теорема о взаимно однозначном соответствии существования полного покрытия и достаточности числа локомотивов.

Ключевые слова: графовая модель, потоки в сетях, назначение локомотивов, грузовые железнодорожные перевозки.

ВВЕДЕНИЕ

Задача составления расписания для грузовых железнодорожных перевозок имеет много различных аспектов и, несмотря на ряд существующих успешных моделей и методов поиска решений, в силу многих специфических особенностей в разных предметных областях остается весьма актуальной. Общий подход к созданию интеллектуальной мультиагентной системы управления перевозочным процессом на железнодорожном транспорте изложен в работах [1, 2]. Описано составление оптимального расписания движения поездов между двумя станциями, соединенными однопутной железной дорогой с разъездом [3, 4]. Рассмотрена задача формирования грузовых составов из имеющихся на станциях заказов и определения маршрутов и расписания их следования по железнодорожной сети так, чтобы минимизировать суммар-

ное взвешенное время выполнения заказов [5]. Задачи назначения локомотивов грузовым составам, оптимизирующие один или несколько критериев, с множеством разнообразных ограничений на функционирование, физические параметры и бюджет, составляют отдельный большой класс задач составления расписаний. Обзор и классификация моделей назначения локомотивов приведены в работе [6]; различные модели и алгоритмы назначения локомотивов даны в работах [7–9].

При описании и формализации предметной области очень важен выбор языка и формального аппарата для моделирования. Для решения задач на двудольных графах применяется многоагентный подход [10, 11]. Предложен метод аукционов, где, помимо многоагентного подхода, применяются стохастические и генетические алгоритмы [11]. В этой же работе впервые сформулирована идея эвристического разбиения временного горизонта планирования на атомарные части, в пределах которых можно предположить соответствие один к одному (локомотив — поезд) и искать глобально-оптимальные (в сумме по атомарному интервалу)

¹ Работа выполнена при частичной финансовой поддержке РФФИ (проект 17-20-01180).

паросочетания методом задачи о назначениях. В результате формируется «временной фронт назначений» распространяющийся по всему горизонту планирования. В работе [12] проблема назначения локомотивов грузовым составам решается поиском локально-оптимальных паросочетаний между двумя списками (локомотивов и поездов) на каждой станции. При этом минимизируется число локомотивов; для поиска решения используются предложенные авторами эвристики. В работе [13] представлен алгоритм решения задачи о назначении локомотивов, основанный на разработанных алгоритмах формирования множества максимальных (по включению) путей ориентированного графа и покрытия вершин ориентированного графа построенным множеством максимальных путей.

В работах [14, 15] решается задача оптимального подбора тяговых ресурсов для грузовых железнодорожных перевозок. Авторами построена модель, цель которой заключается в том, чтобы для заданного графика, состоящего из времен отправления составов с начальных станций и их прибытия на конечные станции (определяемых скоростями движения), а также начального множества локомотивов, задать оптимальный план перевозок — назначить каждому составу локомотив таким образом, чтобы были осуществлены все перевозки и непроизводительное время локомотивов было минимальным. В процессе построения модели на графической (пространственно-временной) плоскости строится альтернативный граф [16], в котором, кроме маршрутов составов, присутствуют также виртуальные ребра нескольких типов: ребра подцепки, перецепки и терминальные ребра. Графы, содержащие ребра разных типов, не все из которых доступны, рассматриваются в моделях графов с нестандартной достижимостью [17, 18]. Потoki, процессы рассеяния, а также сами графовые характеристики (такие, как кратчайшие пути) в этих графах отличаются от обычных. Соответственно, подвергаются модификации и стандартные алгоритмы для их поиска [19, 20]. С этой особенностью сталкивается и описанная далее модель.

Задача, которая будет решаться в настоящей статье, состоит в том, чтобы до поиска оптимального плана назначений, описанного в работе [14, 15], определить, является ли график перевозок выполнимым при заданном числе и начальном расположении локомотивов, и если он невыполним, найти выполнимый подграфик максимальной мощности. Используется идея построения и распространения временного фронта вдоль горизонта планирования, сходная с описанной в работе [11]. Однако помимо отличий в подходах и методах описания сдвига фронта, применяемых в предлагаемой работе, главная ее отличительная черта состоит в смещении цели моделирования. Вместо задачи по-

иска оптимального назначения локомотивов решается более частная проблема: определение максимального подграфика перевозок, для которого заданное множество локомотивов при заданном их начальном расположении является достаточным.

1. ОПИСАНИЕ МОДЕЛИ И ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Рассмотрим участок железной дороги с k станциями.

Множество станций $S = \{S_0, \dots, S_k\}$ линейно упорядоченное.

В модели вводится дискретное время $T = \{1, 2, \dots, T_{fin}\}$. Горизонт планирования T_{fin} составляет, как правило, 24–48 ч.

Множество составов и их траекторий движения $Tr = \{Tr_1, \dots, Tr_n\}$.

Множество локомотивов $L = \{L_1, \dots, L_m\}$.

Начальное положение составов и локомотивов задается отображениями:

$$f_{Tr_start}: Tr \rightarrow (T, S); \quad f_{L_start}: L \rightarrow (T, S).$$

Целевое состояние для составов задается отображением:

$$f_{Tr_goal}: Tr \rightarrow (T, S).$$

Задача перевозки формулируется в терминах пространственно-временного графика.

На графической плоскости (пространственно-временном графике) задан план перевозок грузовых составов на отрезке $[0, T_{fin}]$. На этой же плоскости задано множество локомотивов (рис. 1). Задача состоит в том, чтобы каждому составу назначить, как минимум, один локомотив для его перевозки. Локомотивы могут двигаться порожняком от станции прибытия предыдущего состава к станции отправления следующего. Оптимизация построенных назначений в настоящую задачу не включается.

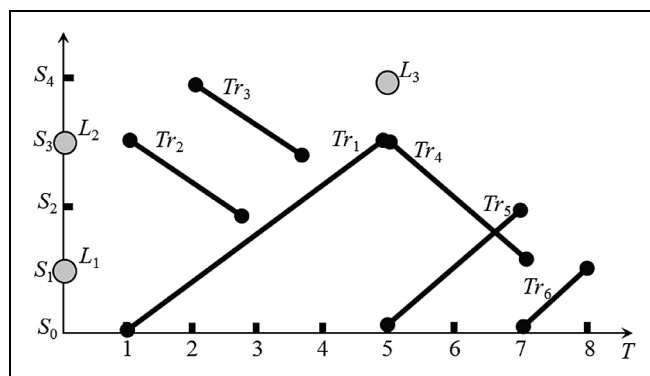


Рис. 1. Графиковая плоскость с заданными перевозками и локомотивами

1.1. Допущения модели

При построении модели примем ряд упрощений.

- Локомотивы характеризуются двумя атрибутами: *допустимой областью тяги* (плечо, домен локомотива) и *грузоподъемностью*. В первой версии модели будем полагать, что все локомотивы и составы находятся в общей области тяги и что грузоподъемности любого локомотива достаточно для любого состава.
- На протяжении всего пути состав не меняет локомотив, т. е. сохраняется соответствие (Tr_i, L_j) .
- Скорости всех локомотивов при осуществлении перевозки составов одинаковы, не зависят от характеристик составов, постоянны и равны v .
- Скорости локомотивов и составов v задаются таким образом, что в них учитывается время подцепки и перецепки локомотива к составу. Таким образом, перецепка локомотива может быть осуществлена в одной точке графической плоскости (t, S_j) . На рис. 1 такая перецепка локомотива возможна между локо-слотами Tr_1 и Tr_4 в момент $t = 5$.
- Поезда не конкурируют при отправлении со станции. В каждый момент времени с одной станции в одном направлении отправляется не более одного поезда. Время между отправлениями поездов всегда не менее заданного интервала безопасности.
- Поезда не конкурируют за пути: все перевозки, представленные на графической плоскости, выполняемы.
- Станции расположены линейно в порядке возрастания номеров. Таким образом:
 - граф железнодорожной сети является простым путем;
 - маршрут между любыми двумя станциями определен однозначно;
 - для расстояний между станциями выполняется равенство

$$|S_{i-1}, S_{i+1}| = |S_{i-1}, S_i| + |S_i, S_{i+1}|.$$

1.2. Основные определения

Определение 1. Отрезки, задающие траектории движения составов на графической плоскости, будем называть *локо-слотами*. ♦

Это название означает, что каждый отрезок Tr_i представляет собой потенциальный слот для подцепки локомотива.

Все локо-слоты имеют одинаковые углы наклона, тангенсы которых равны v и $-v$, где $v = \text{const}$ — скорости локомотивов.

Определение 2. *Графиком G* называется объединение всех локо-слотов на графической плоскости. ♦

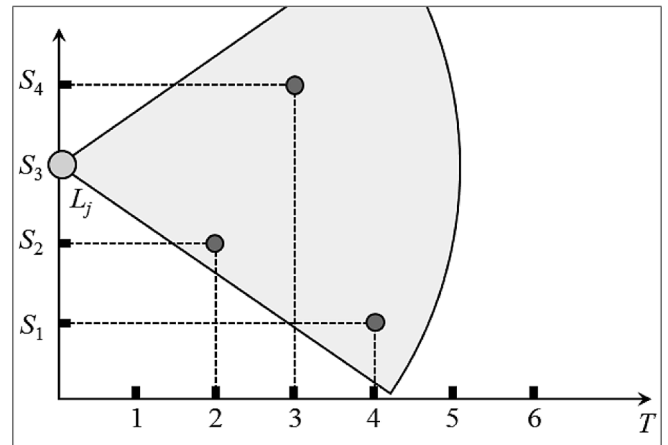


Рис. 2. Конус скорости локомотива L_j

Если локомотива, способного быть подцепленным к локо-слоту Tr_i , не существует, перевозка не может быть выполнена. В этом случае не может быть выполнен и весь график. В данной версии модели будем считать, что график фиксирован и локо-слоты не сдвигаются.

Произвольному локо-слоту Tr_i можно поставить в соответствие все локомотивы, для которых его станция отправления S_{p_i} достижима в момент времени t_i , где $f_{Tr_start}(Tr_i) = (t_i, S_{p_i})$. Определим, какие локо-слоты будут доступны для произвольного локомотива L_j , такого, что $f_{L_start}(L_j) = (t_j, S_{p_j})$.

Определение 3. *Конусом скорости локомотива $Cone(L_j)$* называется пространственно-временной конус на графической плоскости с вершиной в точке $f_{L_start}(L_j) = (t_j, S_{p_j})$, образованный двумя прямыми, тангенс угла наклона которых равен v и $-v$, соответственно. ♦

Для каждого локомотива все доступные локо-слоты находятся в конусе скорости (рис. 2). Заметим, что углы наклонов локо-слотов и направляющих конуса скорости локомотива одинаковы, поэтому, если начало некоторого локо-слота лежит в данном конусе скорости, то в этом конусе лежит весь локо-слот целиком.

Из рис. 2 видно, что конус скорости единственным образом задает времена, в которые данный локомотив может гарантированно оказаться на указанной станции. Так, он может заполнить слоты, начинающиеся в станции S_2 не ранее $t = 2$, S_4 не ранее $t = 3$ и S_1 не ранее $t = 4$.

Определение 4. Все возможные виртуальные ребра, ведущие из локомотивов в начальном состоянии к начальным вершинам локо-слотов, лежащим в их конусах скорости, будем называть *ребрами подцепки*. ♦

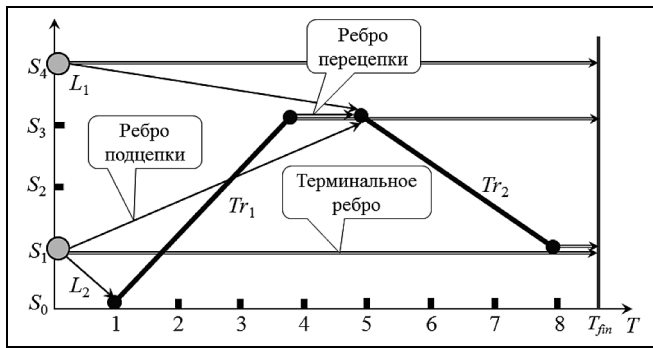


Рис. 3. Ребра подцепки и перецепки и терминальные ребра

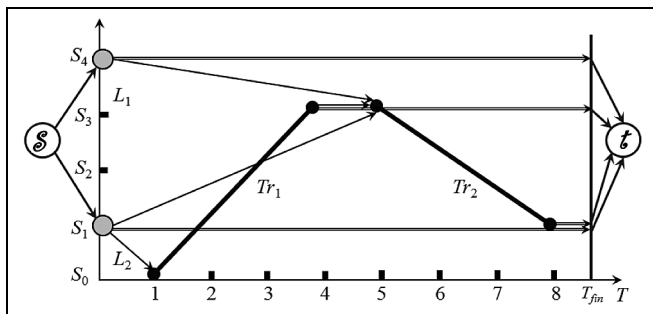


Рис. 4. Источник и сток

За время, составляющее горизонт планирования, как правило, каждый локомотив перевозит более одного состава, поэтому, кроме ребер подцепки локомотивов, существуют также *ребра перецепки*.

Определение 5. Все возможные виртуальные ребра, ведущие из концов одних локо-слотов в начала других, лежащие в их конусах скорости, будем называть *ребрами перецепки*. ♦

Каждый локомотив, закончив свою работу на промежутке времени $[0, T_{fin}]$, движется во времени к горизонту планирования, обозначенному на графике вертикальной линией. Все ребра, ведущие к этой линии от локомотивов и перпендикулярные ей, будем называть *терминальными*. Локомотивы могут переходить в состояние T_{fin} из своего начального состояния (если они не участвовали в перевозках) и из всех концов локо-слотов (рис. 3).

На графике (см. рис. 3) присутствуют четыре терминальных ребра: по одному из начальных позиций локомотивов и из концов двух локо-слотов.

Замечание 1. Тангенсы углов наклона ребер подцепки и перецепки могут принимать любые значения из диапазона $[-v, v]$. На рисунках эти ребра представлены отрезками, в которых средняя скорость локомотива может быть по абсолютной величине меньше v . Такие отрезки — упрощение модели. Они заменяют ломаные линии, состоя-

щие из горизонтальных участков (соответствующих просто локомотивов) и из участков с наклоном, равным $\pm v$, соответствующих движению. ♦

Введем две фиктивные вершины: источник S , из которого выходят связи ко всем локомотивам, и сток t , к которому ведут связи от всех окончаний терминальных ребер (рис. 4).

Таким образом, в графе существуют четыре типа ребер, три из которых виртуальные. Ребра подцепки и перецепки и терминальные ребра будем называть *альтернативными ребрами графа*.

Определение 6. *Простым путем* $Path_i$ называется путь из источника до стока, содержащий альтернативные и терминальные ребра только для локомотива L_i (число локо-слотов в простом пути произвольно). Простой путь, состоящий только из терминального ребра, будем называть *тривиальным*.

Определение 7. *Покрытием* C графика G на альтернативном графе называется такое объединение простых путей $C = \bigcup_{i=1}^m Path_i$, что:

- они попарно не пересекаются $Path_i \cap Path_j = \emptyset \forall i, j \in \overline{1, m}$;
- их объединение содержит весь график G .

Определение 8. График называется *выполнимым*, если для данного множества и начального положения локомотивов существует его покрытие.

Определение 9. Если для данного графика покрытия не существует, график называется *невыполнимым*, а число локомотивов при данном их начальном размещении называется *недостаточным*. ♦

Если число локомотивов недостаточно, будем рассматривать *частичное покрытие* C_{part} графика G , удовлетворяющее первому условию определения покрытия.

Определение 10. Частичное покрытие графика C_{part} будем называть *максимальным по включению* (максимальным), если для данного начального расположения локомотивов *не существует* покрытия C_{part}^* , состоящего из таких простых путей, что

$Path_j \subseteq Path_j^*, j \in \overline{1, m}$, причем, хотя бы один из них содержит одноименный путь из покрытия C_{part} в качестве собственного подмножества: $\exists Path_i^*: Path_i \subset Path_i^*$.

2. СООТНОШЕНИЕ ЧИСЛА ЛОКОМОТИВОВ И ЛОКО-СЛОТОВ

Сформулируем критерий существования полного покрытия графика в терминах числа локомотивов и их начального расположения. Если локомотивов недостаточно для перевозки всех со-

ставов (заполнения всех локо-слотов), то должно быть найдено максимальное по включению покрытие. Если число локомотивов избыточно, возникает проблема отсечения решений, при которых задействуются лишние локомотивы. Эта проблема сводится к задаче редукции множества локомотивов до минимально достаточного. Редуктировать множество локомотивов можно уточнением весов альтернативных ребер — этому будут посвящены наши дальнейшие исследования. В настоящей работе сконцентрируемся на задаче определения достаточности локомотивов и поиске максимального частичного покрытия, если локомотивов недостаточно. Здесь и далее под «максимальным» частичным покрытием будем понимать покрытие, максимальное по включению (см. определение 10).

Понятие недостаточности (как и понятие избыточности) неформально вполне очевидно. Предпримем попытку его формализации.

Пример. Для графика на рис. 5 заведомо не существует покрытия.

В ситуации, представленной на рис. 5, существуют три частичных покрытия, каждое из которых соответствует отдельному терминальному ребру. Два из них — нетривиальны. Ясно, что оптимальным из этих двух частичных покрытий является покрытие, включающее локо-слот Tr_1 ; поскольку станция S_3 , начальная для локомотива, ближе к станции S_4 , чем к станции S_1 (откуда начинаются локо-слоты Tr_1 и Tr_2 соответственно), а прочие условия для двух подцепок одинаковы. ♦

Для того чтобы узнать, достаточно ли число локомотивов для данного графика, воспользовавшись определением, необходимо среди всех возможных комбинаций простых путей найти ту, которая является покрытием графика. Это трудоемкая задача. Если локомотивов недостаточно, требуется полный перебор, чтобы в этом убедиться. Желательно сформулировать более легко проверяемый критерий достаточности локомотивов. Для этого сначала введем дополнительные определения.

Определение 11. Проекцию локо-слота Tr_i на ось времени будем называть *интервалом движе-*

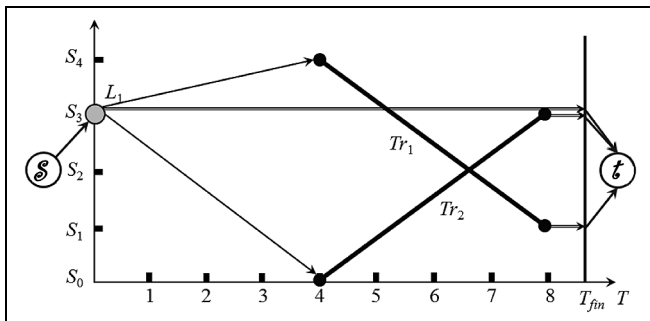


Рис. 5. Два локо-слота с пересекающимися проекциями на шкалу времени и один локомотив

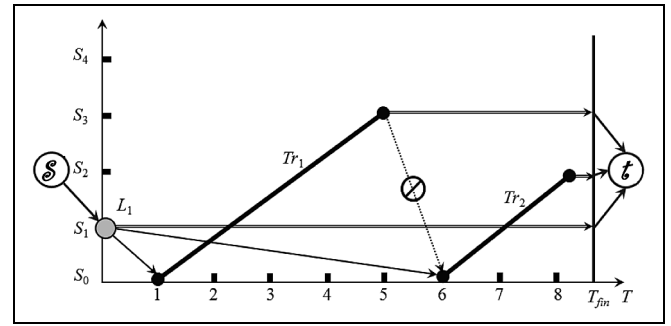


Рис. 6. Недостаточность локомотивов при непересекающихся локо-слотах

ния состава Tr_i и обозначать как I_{Tr_i} : $I_{Tr_i} = [t(f_{Tr_start}(Tr_i)), t(f_{Tr_goal}(Tr_i))]$.

Определение 12. Пересечением локо-слотов будем называть пересечение их интервалов движения: $I_{Tr_i} \cap I_{Tr_j}$.

Определение 13. Мощностью пересечения локо-слотов p_I внутри некоторого интервала $I = [t_1, t_2]$ будем называть максимальное число локо-слотов, пересечение которых непусто и принадлежит интервалу I . ♦

Так, на всем горизонте планирования для графика, представленного на рис. 5, мощность пересечения $p_I = 0$, для графика, представленного на рис. 6, $p_I = 2$.

Утверждение 1 (необходимое условие достаточности локомотивов). Для того чтобы график, представленный объединением локо-слотов, был полным, необходимо, чтобы мощность пересечения локо-слотов на интервале горизонта планирования $I = [0, T_{fin}]$ не превосходила числа локомотивов.

Доказательство. Для существования покрытия графика на интервале максимального пересечения локо-слотов необходимо, чтобы каждому из пересекающихся локо-слотов соответствовал свой локомотив. Если локомотивов меньше, это означает, что покрытия не существует. Тогда локомотивов недостаточно по определению. ♦

Условие утверждения 1 является необходимым для выполнимости графика, однако далеко не всегда достаточным. В нем описывается ситуация, когда несколько поездов должны двигаться одновременно. Тогда, если число локомотивов меньше числа движущихся поездов, их заведомо не хватает.

Вторая ситуация нехватки заключается в том, что интервал между прибытием поезда Tr_i на станцию S_j и отправлением поезда Tr_k со станции S_j меньше, чем время перегона локомотива от станции S_j до станции S_l (плюс время необходимое на отцепку-прицепку локомотива с поездом) (см. рис. 6).

В графике на рис. 6 для локомотива L_1 ребро перецепки, обозначенное пунктиром, не является доступным, поскольку начало локо-слота Tr_2 лежит вне конуса скорости с вершиной в окончании локо-слота Tr_1 .

3. АЛГОРИТМ СДВИГА ФРОНТА

3.1. Описание алгоритма

Для любого фиксированного графика зададим рекурсивный алгоритм построения покрытия, который назовем *алгоритмом сдвига фронта*. Алгоритм, который определяет, достаточно ли число локомотивов, должен строить покрытие всегда, когда оно существует. Если покрытие может быть построено, локомотивов достаточно.

В графиках примем такую нумерацию локо-слотов, что $i < j$ для Tr_i и Tr_j , если и только если $t(f_{Tr_start}(Tr_i)) \leq t(f_{Tr_start}(Tr_j))$. Таким образом, локо-слоты упорядочены по времени отправки поездов.

Определение 14. *Фронт* — это виртуальная ломаная линия, соединяющая начала l локо-слотов с самыми маленькими номерами, достижимых из точек старта l локомотивов, $l \leq m$, где $m = |L|$ — число локомотивов. ♦

Иными словами, фронт — это те выбранные локо-слоты, каждому из которых к моменту его начала можно поставить во взаимно однозначное соответствие доступный локомотив.

Фронт распространяется слева направо таким образом, чтобы по возможности ни один локо-слот не был пропущен. Точки старта локомотивов могут быть распределены не только в пространстве, но и во времени. (Например, локомотив L_3 — см. на рис. 1 — становится доступным только в момент времени $t = 5$.) Локомотив включается в алгоритм, когда он впервые оказывается позади (левее) идущего фронта. Если какой-то локо-слот после прохождения фронта через его начало остался незаполненным, это означает, что покрытия

данного графика не существует (пока это часть определения — для конкретного алгоритма построения фронта это утверждение будет нуждаться в доказательстве). При заполнении (и прохождении) локо-слотов, формирующих фронт с номером k , локомотивы попадают в новые точки графической плоскости, и их положение определяет фронт $k + 1$.

Рассмотрим алгоритм сдвига фронта на примере. Пусть график представляет собой локо-слоты, заданные на рис. 7. Пусть, кроме этого, имеются три локомотива, стартующих в нулевой момент времени.

Шаги алгоритма. Все локо-слоты, начала которых лежат в конусе скорости локомотива, соединим с данным локомотивом ребрами подцепки. Для графика (рис. 7) соответствия можно задать таблицей, в ячейках которой перечислены номера доступных локо-слотов:

Локомотив	Фронт	
	0	...
L_1	1, 4, 5, 6	
L_2	3, 4, 5, 6	
L_3	2, 3, 4, 5, 6	

Поскольку локомотивов три, может быть подцеплено не более трех составов. Для формирования фронта нужно оставить три локо-слота с наименьшими номерами (в данном случае, это номера 1, 2, 3), получим все возможные комбинации подцепки для фронта 0:

Локомотив	Фронт	
	0	...
L_1	1	
L_2	3	
L_3	2, 3	

Отсюда сразу видно, что существует единственная подцепка, при которой будут покрыты все три локо-слота (см. рис. 8).

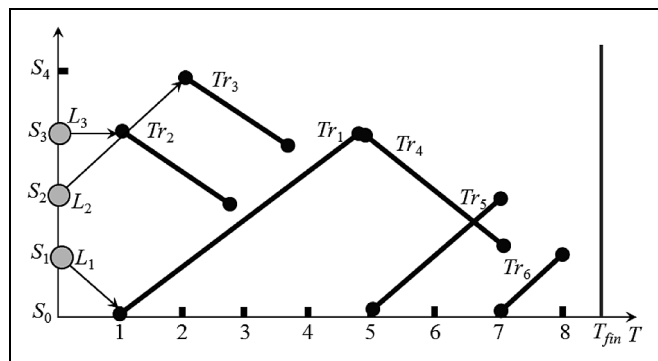


Рис. 7. График и локомотивы

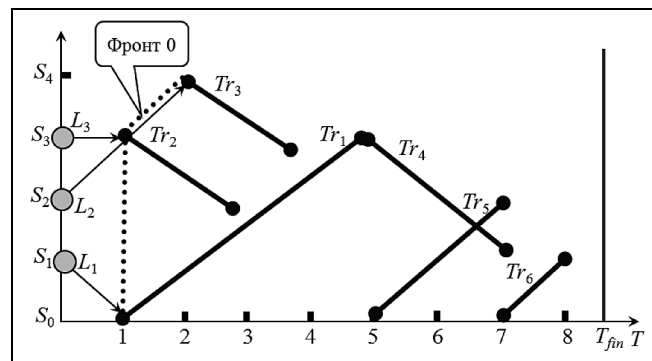


Рис. 8. Фронт 0

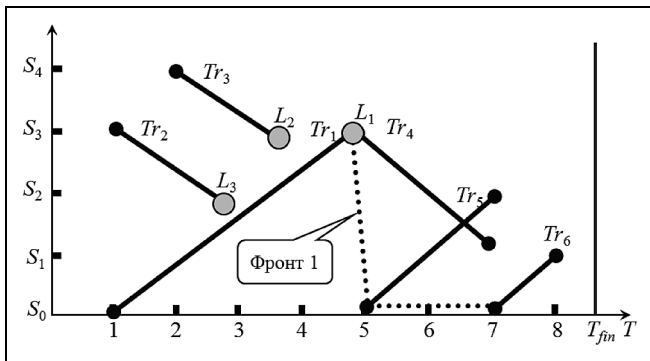


Рис. 9. Новый фронт (фронт 1)

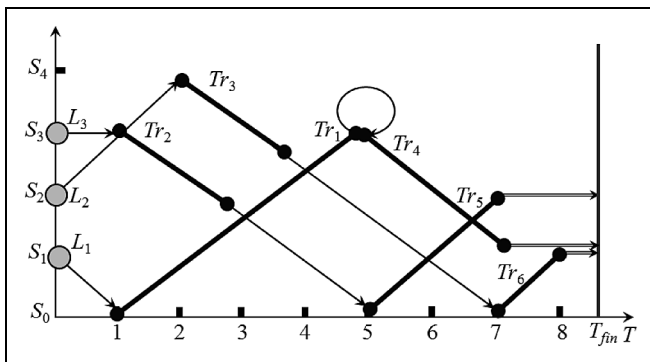


Рис. 10. Покрытие графика

В общем виде, если описывающие фронт ячейки содержат по несколько значений, задачу о поиске соответствий между локомотивами и номерами локо-слотов можно представить в виде двудольного графа. В таком виде решение сводится к поиску наибольшего паросочетания [21].

Замечание 2. Кажется бы, можно сразу построить ребра подцепки только к первым m локо-слотам, где m — число локомотивов, чтобы затем не отбрасывать лишнее. Однако это можно делать только в том случае, когда все m локо-слотов доступны и существует, по крайней мере, одно взаимно однозначное соответствие между локомотивами и локо-слотами. Если наибольшее паросочетание не покрывает m первых локо-слотов, график при таких начальных данных не является выполнимым. Однако алгоритм может продолжить работу для построения частичного покрытия. ♦

Положения локомотивов в концах локо-слотов фронта 0 определяют фронт 1. Новый фронт представляет собой начала локо-слотов с номерами 4, 5 и 6 (рис. 9). Напомним, что время перецепки в случае локо-слотов учитывается в скоростях, и перецепка $Tr_1 \rightarrow Tr_4$ является возможной.

Считая положение локомотивов (см. рис. 9) новым начальным состоянием, повторяем первый

шаг алгоритма. Нетрудно проверить, что ребра подцепки для фронта 1 будут такими:

Локомотив	Фронт		
	0	1	...
L_1	1	4	
L_2	3	6	
L_3	2	5, 6	

Как и прежде, видно, что распределение локомотивов по локо-слотам, покрывающее все три слота, единственно. Имеем два фронта, с помощью которых задается покрытие всего графика (рис. 10):

Локомотив	Фронт	
	0	1
L_1	1	4
L_2	3	6
L_3	2	5

Замечание 3. Как видно из описания алгоритма, он устойчив к появлению или выведению локомотивов на протяжении всего горизонта планирования. Добавленные локомотивы будут формировать ближайший справа к моменту добавления фронт. Выведенные локомотивы не будут принимать участие в формировании нового фронта. Никаких дополнительных изменений вносить не потребуется.

3.2. Условия применимости алгоритма сдвига фронта

Достижение фронтом горизонта планирования с покрытием всех локо-слотов является *достаточным условием достаточности числа локомотивов*.

Однако возможны ситуации, в которых описанный алгоритм не строит *покрытие*, хотя оно существует. Рассмотрим пример, представленный на рис. 11.

Применяя предложенный алгоритм, на фронте 0 получим:

Локомотив	Фронт	
	0	1
L_1	1, 2, 3	...
L_2	2, 3	...

Видно, что на фронте 0 достижимы оба первых локо-слота. Тогда:

Локомотив	Фронт	
	0	...
L_1	1	...
L_2	2	...

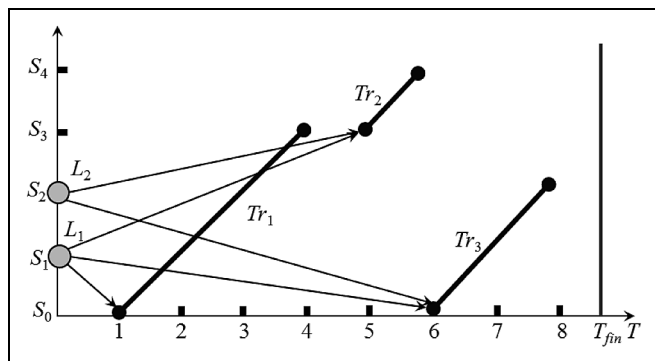


Рис. 11. Пример графика, для которого алгоритм сдвига фронта не строит покрытие

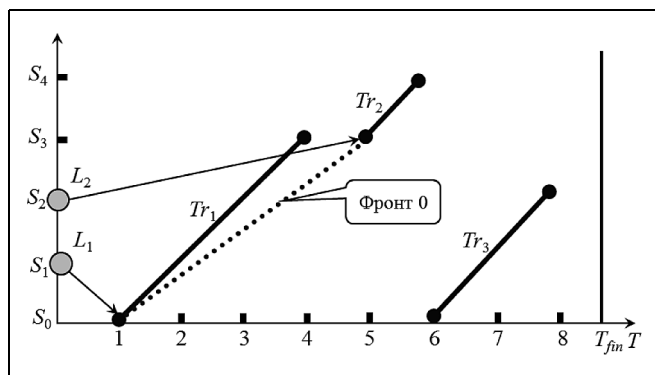


Рис. 12. Фронт 0

Из рис. 11 и 12 видно, что из концов первых двух локо-слотов нельзя попасть в начало локо-слота 3. Таким образом, локо-слот 3 покрыть не удастся. Фронт 1 и покрытие целиком не существуют:

Локомотив	Фронт	
	0	1
L_1	1	⊗
L_2	2	⊗

Однако легко заметить, что если локо-слоты 1 и 2 оставить для первого локомотива, то для этого графика существует покрытие:

Локомотив	Фронт	
	0	1
L_1	1	2
L_2	3	⊗

Ясно, что алгоритм сдвига фронта в описанном выше виде не может построить такое покрытие, потому что локо-слоты в нем распределены по фрон-

там не в порядке возрастания номеров. В этом покрытии утрачена монотонность (фронту с большим номером должны соответствовать локо-слоты с большими номерами).

Для того чтобы алгоритм работал корректно и всегда находил покрытие, если оно существует, нужно произвести *нормализацию фронта*.

3.3. Процедура нормализации фронта

Рассмотрим на примере графика (см. рис. 11 и 12), в чем заключается проблема описанного алгоритма. Из рис. 12 видно, что конец локо-слота Tr_1 оказался позади фронта 0. Это означает, что на начало формирования фронта 1 локомотив L_1 будет по-прежнему находиться в зоне действия фронта 0. Именно из-за этого произошла коллизия с распределением локо-слотов для локомотивов. Чтобы этого избежать, нужно сделать так, чтобы каждый локомотив при переходе в конец локо-слота, занятого на фронте k , проходил через линию фронта k . Таким образом, у каждого локо-слота, принадлежащего фронту k , конец должен лежать правее этого фронта.

При каком условии локо-слот Tr_j не переходит через фронт k ? Это случается, когда в этом же фронте существует другой локо-слот Tr_p , который начинается после окончания локо-слота Tr_j . Более формально, нормализация имеет смысл, если локо-слот Tr_i оказался в конусе скорости, вершина которого находится в конце локо-слота Tr_j . Иначе Tr_i не будет достижим из Tr_j , и поэтому его принадлежность текущему и следующему фронту даст один и тот же результат. Если на рис. 11 и 12 начало локо-слота Tr_2 сместить вниз на станцию S_0 , график не будет выполнен ни при каком распределении слотов по фронтам.

Нормализация фронта будет заключаться в том, что все локо-слоты Tr_p , начало которых лежит правее локо-слота Tr_j в конусе скорости, построенном из его конечной точки, удаляются из фронта k и переходят сразу на фронт $k + 1$. Тогда на фронте k будут задействованы не все локомотивы, а только те, которые соответствуют локо-слотам, недостижимым из конца Tr_p . Это позволит сохранить монотонность номеров локо-слотов — их возрастание в каждом последующем фронте.

Для ситуации, представленной на рис. 11 и 12, процедура нормализации будет выглядеть следующим образом. Из рис. 12 видно, что на фронте 0 локо-слот Tr_2 попадает в конус скорости из окончания локо-слота Tr_1 . Поэтому локо-слот Tr_2 исключается из фронта 0. Тогда локомотив L_2 тоже элиминируется из участия во фронте 0, так как на этом фронте для него нет подходящего слота.

Фронт 0 состоит только из одной точки — начала локо-слота Tr_1 (рис. 13):

Локомотив	Фронт	
	0	1
L_1	1	...
L_2	⊗	...

Фронт 1 в этом случае будет выглядеть, как на рис. 14:

Локомотив	Фронт	
	0	1
L_1	1	2
L_2	⊗	3

Как видно из примера, при нормализации фронта монотонность восстанавливается.

Модифицированный алгоритм назовем *алгоритмом сдвига фронта с нормализацией*.

Замечание 4. Легко увидеть, что по столбцам таблиц алгоритма записаны фронты, а по строкам — простые пути $Path_i$. Причем, заведомо (по построению) выполняется первое свойство покрытия графика G : пути попарно не пересекаются $Path_i \cap Path_j = \emptyset \forall i, j \in \overline{1, m}$. Если выполняется

второе свойство: «их объединение содержит весь график G », то строки таблицы после выполнения алгоритма содержат покрытие. Если второе свойство не выполняется, то с помощью алгоритма было найдено лишь частичное покрытие. ♦

Докажем, что если алгоритм не смог найти полное покрытие графика, то такого покрытия не существует.

Для этого сначала докажем вспомогательные утверждения.

Лемма 1. Если при выполнении алгоритма сдвига фронта с нормализацией на графике G на фронте с номером k выбрано $l \leq t$ локо-слотов и существует несколько способов распределить их между l локомотивами, все назначения локомотивов эквивалентны относительно существования покрытия.

Доказательство. Рассмотрим положения локомотивов после прохождения фронта $k - 1$ или начальные положения, если $k = 0$. Если начало локо-слота Tr_i , принадлежащего фронту k , лежит в пересечении конусов скорости нескольких локомотивов, любой назначенный локомотив окажется после прохождения фронта в конце локо-слота Tr_i , откуда будет строиться новый конус скорости. При разных назначениях получатся разные простые пути для каждого локомотива в отдельности. Однако их объединение будет содержать одни и те же локо-слоты в силу построения. А, следовательно, существование покрытия инвариантно относительно альтернативных назначений локомотивов.

Лемма 2. Для данного графика G при данном множестве и начальном расположении локомотивов

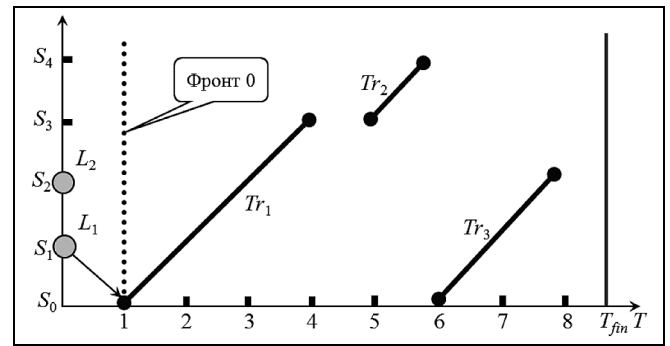


Рис. 13. Фронт 0 с участием одного из двух локомотивов

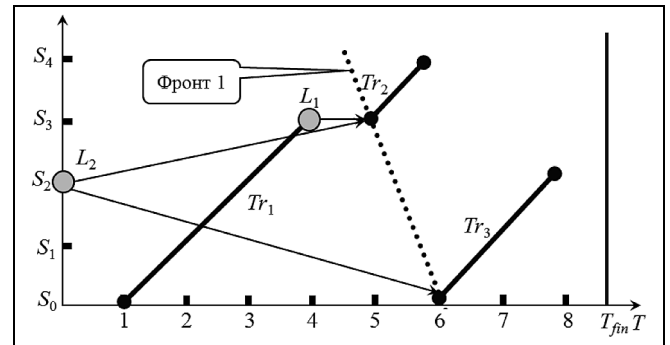


Рис. 14. Фронт 1, формирующий покрытие

нормализованные фронты задаются единственным образом.

Доказательство. Это утверждение верно по построению алгоритма. Не единственность возможна только в одном случае: если на фронте k существует более одного локо-слота с началами, лежащими в конусе скорости локомотива L_j (и только L_j); т. е. выбор состоит в том, какие из локо-слотов останутся не заполненными. В этом случае покрытия не существует. Для построения частичного покрытия добавим в алгоритм еще одно условие: будем считать, что в случае альтернативного выбора заполнения одного из нескольких локо-слотов выбирается локо-слот с наименьшим номером.

Теорема 1. Покрытие графика G при заданном начальном числе и положении локомотивов существует тогда и только тогда, когда оно может быть построено с помощью применения алгоритма сдвига фронта с нормализацией.

Доказательство. Достаточность в утверждении теоремы очевидна: если покрытие построено, то оно существует.

Необходимость докажем от противного. Пусть существует покрытие графика G . И пусть результатом алгоритма сдвига фронта с нормализацией является частичное покрытие графика G . По лемме 1, если алгоритм разветвляется, все его ветви эквивалентны относительно существования покрытия. Это означает, что если построено частичное покрытие, никакое другое назначение локомотивов с помощью алгоритма сдвига с нормализацией не даст полного покрытия. Тогда имеется

по крайней мере один локо-слот Tr_j , не попавший в таблицу алгоритма. По лемме 2 каждый фронт строится единственным образом.

Рассмотрим две ситуации.

1. В таблице алгоритма существует столбец k , содержащий номера слотов, меньшие j , и, по крайней мере, один из номеров, больших j . Это означает, что локо-слот Tr_j должен был оказаться во фронте k , поскольку алгоритм учитывает нормализацию фронта, номера локо-слотов распределяются по фронтам монотонно. Фронт k содержит не более чем m локо-слотов. Из фронта $k - 1$ и, возможно, более ранних при наличии нормализации к нему ведет столько же ребер подцепки. Если $k = 0$, ребра подцепки идут непосредственно из начального положения локомотивов. Слот Tr_j элиминируется из фронта k , если и только если все ребра подцепки, доступные для Tr_j , заняты слотами с меньшими номерами. А ребра, ведущие к следующим по номерам локо-слотам Tr_l $l > j$, из фронта k , не могут быть переброшены на локо-слот Tr_j . Ни при какой другой расстановке локомотивов не существует возможности поместить на фронте k локомотивы с номерами без пропусков. Тогда, поскольку к фронту k ведут ребра подцепки из ближайших m локомотивов, то нельзя построить покрытие, содержащее все слоты, принадлежащие фронту k , и слот Tr_j . Отсюда следует, что покрытия графика не существует.

2. В таблице алгоритма существует столбец $k - 1$, содержащий номера слотов, меньшие j , и столбец k , все номера слотов в котором больше j . Это означает, что локо-слот Tr_j уже не вошел в фронт $k - 1$, но еще не вошел в фронт k . Рассмотрим фронт $k - 1$. Если в столбце $k - 1$ заняты все m строк, то номер j должен принадлежать столбцу k . Если в нем занято меньше m строк, то была произведена нормализация фронта, и номер j снова должен принадлежать столбцу k . В обоих случаях мы попадаем в ситуацию из п. 1, и покрытия графика не существует, что противоречит условию теоремы.

Что и требовалось доказать.

Следствие. Алгоритм сдвига фронта с нормализацией строит максимальное по включению частичное покрытие графика G при заданном начальном числе и положении локомотивов.

3.4. Оценка сложности алгоритма сдвига фронта с нормализацией

Оценим алгоритмическую сложность нахождения покрытия.

- Для определения фронта k требуется построение не более чем $C_1 m$ ребер подцепки.
- Нормализация фронта k требуется только в том случае, если конец одного из не более чем m локо-слотов лежит левее начала другого локо-слота из этого же множества. Поскольку локо-слоты упорядочены, проверка должна начинаться с правого конца. Если такой локо-слот найден, нужно проверить, попадает ли, по крайней мере, один локо-слот из фронта k , лежащий правее, в его конус скорости. Сложность такой процедуры не более чем $C_2 m^2$.

- Для назначения локомотивов локо-слотам на текущем фронте требуется найти наибольшее паросочетание в двудольном графе, в одной доле которого локомотивы, а в другой — локо-слоты. Задача поиска наибольшего паросочетания в двудольном графе является одной из алгоритмически простых задач, она решается многими методами, в том числе и потоковыми [22], при которых строятся фиктивный источник и сток. Алгоритм Хопкрофта — Карпа для двудольного графа имеет сложность, как максимум, $O(EV^{1/2})$. В нашем случае $E < m^2$, $V \leq 2m$. Если $E = m^2$, сочетания создаются произвольно за один шаг.

- Число фронтов $\sim n/m$.

Оценивая суммарное значение, получим: $(C_1 m)(C_2 m^2)(C_3 m^{3/2})(n/m) = C \cdot n \cdot m^{7/2}$. Общая сложность алгоритма, таким образом, составит в худшем случае $O(nm^{7/2})$. Однако, в общем случае, эта оценка сильно завышена.

ЗАКЛЮЧЕНИЕ

Предложен алгоритм, позволяющий однозначно определить, достаточно ли число локомотивов для выполнения данного графика. Результат его работы так называемое покрытие графика или максимальное частичное покрытие, если полного не существует. Покрытие — способ осуществить все перевозки средствами данных локомотивов; оно представляет собой объединение простых путей. Простой путь — путь от некоторого локомотива до стоковой вершины. В результате работы алгоритма получается столько путей, сколько локомотивов в системе. Каждый путь проходит через несколько фронтов, зависящих от числа и начального расположения локомотивов. В этом смысле алгоритм сдвига фронта сходен с алгоритмами перколяции [23—26]. Теория перколяции возникла в физике как теория, исследующая задачи просачивания жидкостей в пористой среде, а также некоторые задачи, связанные с проводимостью. Формально такого рода задачи можно рассматривать на графе, вершинам которого соответствуют пустоты пористой среды, а ребрам — каналы, соединяющие пустоты. Процесс просачивания выглядит при этом как постепенное и случайное заполнение вершин графа, при котором в графе возникают пути. Особое внимание в теории перколяции уделяется регулярным графам. На конечных решетках исследуют появление такого состояния, при котором существует хотя бы один непрерывный путь одного края решетки до противоположного. В случае алгоритма сдвига фронта случайные процессы заменяются детерминированными, при этом перколяция происходит через фронты и заканчивается образованием m путей, прошедших через весь граф.



Таким образом, видно, что при существовании нескольких полных покрытий графика, помимо стандартных методов поиска оптимального покрытия, можно применять неклассические методы, приводящие к решению задачи оптимизации на построенном графе. Выбор этих методов и нахождение оптимального решения для произвольного графика, так же, как описание свойств полученных графов — предмет дальнейших исследований.

ЛИТЕРАТУРА

1. Шабунин А.Б., Чехов А.В., Сазуров С.В. и др. Разработка сетцентрического подхода к созданию интеграционной платформы и интеллектуальной системы управления ресурсами грузовых сортировочных станций в реальном времени // Тр. третьей российской конф. с междунар. участием «Технические и программные средства систем управления, контроля и измерения» (УКИ'12), Москва, 16–19 апреля 2012, ИПУ РАН. — М., 2012. — С. 1560–1572.
2. Шабунин А.Б., Марков С.Н., Дмитриев Д.В. и др. Интеграционная платформа для реализации сетцентрического подхода к созданию распределенных интеллектуальных систем управления ресурсами ОАО «РЖД» // Программная инженерия. — 2012. — № 9. — С. 23–28.
3. Лазарев А.А., Мусатова Е.Г., Тарасов И.А. Решение задачи планирования двухстороннего движения на однопутном участке железной дороги с разъездом // Автоматика и телемеханика. — 2016. — № 11. — С. 158–174.
4. Gafarov E., Dolgui A., Lazarev A. Two-Station Single-Track Railway Scheduling Problem With Trains of Equal Speed // Comput. and Industrial Engin. — 2015. — Vol. 85. — P. 260–267.
5. Лазарев А.А., Мусатова Е.Г. Целочисленные постановки задачи формирования железнодорожных составов и расписания их движения // Управление большими системами. — 2012. — Вып. 38. — С. 161–169.
6. Piu F. and Speranza M.G. The locomotive assignment problem: a survey on optimization models // Intl. Trans. in Op. Res. — Vol. 21. — P. 327–352. DOI: 10.1111/itor.12062.
7. Ahuja R.K., Liu J., Orlin J.B., et al. Solving real-life locomotive scheduling problems // Transportation Science. — 2005. — Vol. 39, N 4. — P. 503–517.
8. Jaumard B., Tian H. Multi-Column Generation Model for the Locomotive Assignment Problem // Proc. of 16th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS'16). — 2016. — P. 6:1–6:13.
9. Kasalica S., Mandić D., Vukadinović V. Locomotive Assignment Optimization Including Train Delays // Promet — Traffic & Transportation. — 2013. — Vol. 25, N 5. — P. 421–429.
10. Такмазян А.К., Шелудяков А.В. Мультиагентное решение методом аукционов многопродуктовой транспортной задачи с объединенными потребностями // Тр. четвертой науч.-техн. конф. с междунар. участием «Интеллектуальные системы управления на железнодорожном транспорте. Компьютерное и математическое моделирование», ИСУЖТ-2015. — М., 2015. — Сер. 1. — С. 110–112.
11. Кузнецов Н.А., Пащенко Ф.Ф., Рябых Н.Г. и др. Алгоритмы оптимизации в задачах планирования на рельсовом транспорте // Информационные процессы. — 2014. — Т. 14, № 4. — С. 307–318.
12. Azanov V.M., Buyanov M.V., Gaynanov D.N., Ivanov S.V. Algorithm and software development to allocate locomotives for transportation of freight trains // Вестник ЮУрГУ. Сер. Матем. моделирование и программирование. — 2016. — Т. 9, № 4. — С. 73–85.
13. Гайнанов Д.Н., Кибзун А.И., Рассказова В.А. Теоретико-графовый алгоритм решения задачи о назначении и перемещении локомотивов // Вестник компьютерных и информационных технологий. — 2017. — № 5. — С. 51–56.
14. Matyukhin V.G., Shabunin A.B., Kuznetsov N.A., Takmazjan A.K. Rail Transport Control by Combinatorial Optimization Approach. 11th IEEE International Conference on Application of Information and Communication Technologies AICT2017. Conf. Proc., IEEE Communications Society. — 2017. — Vol. 1. — P. 419–422.
15. Матюхин В.Г., Шабунин А.Б., Кузнецов Н.А. и др. Графовая динамическая модель задачи подбора тяговых ресурсов для грузовых железнодорожных перевозок // Тр. шестой науч.-техн. конф. «Интеллектуальные системы управления на железнодорожном транспорте. Компьютерное и математическое моделирование» (ИСУЖТ-2017), АО «НИИАС». — М., 2017. — С. 14–18.
16. Mascis A., Pacciarelli D. Job-shop scheduling with blocking and no-wait constraints. European Journal of Operations Research. — 2005. — Vol. 143, N 3. — P. 498–517.
17. Ерусалимский Я.М. Потоки в сетях с нестандартной достижимостью // Известия вузов. Северо-Кавказский регион. Ест. науки. — 2012. — № 1. — С. 5–7.
18. Ерусалимский Я.М., Скороходов В.А., Кузьмина М.В., Петросян А.Г. Графы с нестандартной достижимостью: задачи, приложения: монография. — Ростов-на-Дону: ЮФУ, 2009. — 195 с.
19. Ловас Л., Пламмер М. Прикладные задачи теории графов. Теория паросочетаний в математике, физике, химии: пер. с англ. — М.: Мир, 1998. — 653 с.
20. Седжевик Р. Фундаментальные алгоритмы на C++. Алгоритмы на графах. — СПб.: ДиаСофтЮП, 2002. — 496 с.
21. Lovász L., Plummer M.D. Matching Theory. — North-Holland, 1986. — 554 p.
22. Ahuja R.K., Magnanti T.L., Orlin J.B. Network Flows: Theory, Algorithms, and Applications. — Prentice Hall, United States, 1993.
23. Bollobas B., Riordan O. Percolation: 2nd ed. — Cambridge: Cambridge Univ. Press, 2009.
24. Тарасенко Ю.Ю. Перколяция: Теория, приложения, алгоритмы: 2-е изд. — М.: Либроком, 2012.
25. Кузнецов О.П. Сложные сети и распространение активности // Автоматика и телемеханика. — 2015. — № 12. — С. 3–26.
26. Radicchi F. Percolation in real interdependent networks // Nature Physics. — 2015. — Vol. 11. — P. 597–602. DOI:10.1038/nphys3374.

Статья представлена к публикации членом редколлегии А.А. Лазаревым.

Жиликова Людмила Юрьевна — д-р физ.-мат. наук, вед. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, zhilyakova.ludmila@gmail.com,

Кузнецов Николай Александрович — академик РАН, советник РАН; зав. лабораторией, Институт радиотехники и электроники им. В.А. Котельникова РАН, г. Москва, kuznetsov@cplire.ru,

Матюхин Владимир Георгиевич — д-р техн. наук, первый зам. ген. директора, Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте, г. Москва,

Шабунин Александр Борисович — руководитель науч.-техн. комплекса, Научно-исследовательский и проектно-конструкторский институт информатизации, автоматизации и связи на железнодорожном транспорте, г. Москва, a.shabunin@vniias.ru,

Такмазян Андрей Куркенович — канд. физ.-мат. наук, гл. специалист, ООО «Програмпарк», г. Москва, takmazjan@gmail.com.



ДЕСЯТАЯ МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ «УПРАВЛЕНИЕ РАЗВИТИЕМ КРУПНОМАСШТАБНЫХ СИСТЕМ» MLSD'2017

Конференция состоялась 2—4 октября 2017 г. в Институте проблем управления им. В.А. Трапезникова РАН и продолжила научную работу в интересах развития теории и практики управления развитием крупномасштабных систем.

Конференции по вопросам управления развитием крупномасштабных систем проводятся ежегодно, начиная с 2007 г., и накопили большой опыт взаимодействия с научной и научно-промышленной общественностью страны и ближнего зарубежья по актуальным проблемам управления развитием крупномасштабных систем, плодотворного обмена опытом и совместных разработок в вопросах долгосрочного планирования и стратегического развития крупномасштабных систем, исследований в области научно-методического сопровождения новых разработок в сфере развития крупномасштабных систем.

Тематика конференций развивалась и расширялась на протяжении последнего десятилетия и в настоящий момент представляет широкий спектр проблем от управления развитием топливно-энергетических, транспортных, инфраструктурных, организационно-технических комплексов и систем, крупномасштабных производств и технологий до управления объектами повышенной опасности, мониторинга в задачах управления крупномасштабными системами, систем здравоохранения, медико-биологических систем и технологий, систем обработки и интеллектуального анализа больших массивов информации.

На MLSD'2017 работали 13 секций по следующим направлениям.

- Проблемы управления развитием крупномасштабных систем, включая ТНК, госхолдинги и госкорпорации.
- Методы и инструментальные средства управления инвестиционными проектами и программами.

- Проектные офисы — институты развития крупномасштабных систем.
- Имитация и оптимизация в задачах управления развитием крупномасштабных систем.
- Управление топливно-энергетическими, инфраструктурными и другими системами.
- Управление транспортными системами.
- Управление развитием авиационно-космических и других крупномасштабных организационно-технических комплексов и систем.
- Управление региональными, городскими, муниципальными системами.
- Управление объектами атомной энергетики и другими объектами повышенной опасности.
- Информационное и программное обеспечение систем управления крупномасштабными производствами.
- Мониторинг в задачах управления крупномасштабными системами.
- Управление развитием крупномасштабных систем здравоохранения, медико-биологических систем и технологий.
- Методология, методы и программно-алгоритмическое обеспечение обработки и интеллектуального анализа больших массивов информации.

В программу MLSD'2017 были включены 24 пленарных и 256 секционных докладов, представленных учреждениями Российской академии наук, вузами, академиями и университетами, а также управленческими и коммерческими организациями и рядом зарубежных научных учреждений. В работе конференции участвовало более 450 чел. Значительное число участников составляли молодые ученые и специалисты. Было представлено несколько десятков докладов по результатам работ, выполненных при финансовой поддержке РФФИ и других российских фондов.

Работу конференции открыл директор Института, чл.-корр. РАН Д.А. Новиков, отметивший расту-



ший интерес к проблемам развития теории и практики управления крупномасштабными системами.

В базовом докладе академика *С.Н. Васильева* и д-ра техн. наук *А.Д. Цвиркуна* упор сделан на проблемы разработки методологии и инструментальных средств управления развитием крупномасштабных систем для решения стратегических задач управления социально-экономическим, промышленно-территориальным развитием России.

Большое внимание в своем выступлении докладчик *А.Д. Цвиркун* уделил проблемам управления, связанным с особенностями крупномасштабных систем, и главной задаче — созданию методологии разработки стратегий развития, опирающейся на институциональные реформы и согласование различных видов социально-экономической политики, и разработке стратегии развития на основе предложенной методологии. На междисциплинарной основе необходимо разработать инструменты и механизмы выбора оптимальных решений, мер государственного регулирования, управления и государственно-частного партнерства, обеспечивающие опережающий экономический рост и промышленно-инфраструктурную модернизацию страны.

Доклад академика *С.П. Филиппова* «Управление научно-технологическим развитием энергетики» был посвящен вопросам прогнозирования научно-технического прогресса в энергетике, анализу новых энергетических технологий, теоретическим и прикладным исследованиям в задачах выбора приоритетных технологий и обоснования научно-исследовательских и конструкторских работ в области тепловой и возобновляемой энергетики, глубокой переработки топлив, теплоснабжения, энергосбережения.

Докладчик поделился опытом прогнозных системно-технологических исследований в энергетике, глобальных энергетических исследований, прогнозирования развития топливно-энергетического комплекса страны и регионов на долгосрочную перспективу, разработки системы взаимосогласованных топливно-энергетических балансов, прогнозирования энергопотребления страны и регионов, исследования проблем энергоэффективности, в том числе технико-экономического анализа энергосберегающих технологий, и др.

Тему развития энергетического комплекса страны продолжил доклад чл.-корр. РАН *Н.И. Вороня* и канд. экон. наук *Г.И. Шевелевой* «Особенности и перспективы развития корпоративного управления в российской электроэнергетике», представленный *Г.И. Шевелевой*. В целях совершенствования корпоративного управления в российских генерирующих компаниях электроэнергетики коллек-

тивом исследователей из Института систем энергетики им. Л.А. Мелентьева СО РАН выполнен углубленный анализ действующей корпоративной практики энергетических компаний страны, в том числе в сравнении с зарубежной практикой. По результатам анализа выявлены особенности корпоративного управления в российских генерирующих компаниях электроэнергетики и приоритетные направления его дальнейшего развития, в том числе необходимость повышения профессиональных компетенций руководства и членов советов директоров корпораций в вопросах экономической, правовой и организационно-управленческой культуры.

Проблемы энергетического развития в сфере разработки, добычи и переработки углеводородного сырья были рассмотрены в докладах чл.-корр. РАН *В.И. Салыгина* и канд. экон. наук *И.А. Гулиева* «Современные вызовы добычи и транспортировки углеводородов в Каспийском регионе», канд. техн. наук *А.В. Ахметзянова*, д-ра физ.-мат. наук *А.Г. Кушнера*, д-ра физ.-мат. наук *В.В. Лычагина* «Модель фильтрации Бакли — Леверетта при оптимальном управлении разработкой нефтяных месторождений», д-ра техн. наук *Л.Р. Соркина*, канд. техн. наук *Ю.Р. Шишорина*, канд. техн. наук *Ю.М. Цодикова*, канд. техн. наук *Н.В. Мостового* «Разработка и моделирование программ долгосрочного развития предприятий нефтепереработки и нефтехимии».

Доклад *В.И. Салыгина* и *И.А. Гулиева* касался разрешения вопросов международного сотрудничества в использовании природных ресурсов Каспийского моря пятью сопредельными государствами, главным образом, в нефтегазовой деятельности, которое невозможно без обеспечения безопасности окружающей среды. Перед государствами региона стоит задача обеспечения взаимовыгодного и рационального использования всех ресурсов Каспийского моря. Такая задача решается путем многостороннего международного сотрудничества. На данный момент страны Каспийского региона заключили ряд соглашений, определяющих статус Каспийского моря и регулирующих отношения по совместной разработке и эксплуатации месторождений.

В докладе, представленном *Ю.Р. Шишориным*, изложен опыт разработки программ развития и разработки технологических моделей вариантов для этапов реконструкции нескольких заводов при решении задачи модернизации нефтеперерабатывающих и нефтехимических предприятий. Выбор оптимального варианта технологического развития НПЗ/НХК на 5—10 лет и более осуществляется с учетом многокритериального финансово-экономического анализа. Основой для такого комп-

лексного исследования служат оптимизационные модели, как правило, нелинейные, и соответствующие им задачи решаются методом последовательного линейного программирования. Моделирование реконструкции и развития НПЗ включает в себя моделирование вариантов технологической схемы предприятия для каждого этапа программы развития и моделирование инвестиционного процесса с учетом всех затрат, движения денежных средств, получения кредитов, времени на реконструкцию и строительство. Разработанная методика успешно применена для реализации программ развития нескольких крупных производственных комплексов.

Доклад д-ра техн. наук *В.В. Баранова* и д-ра техн. наук *А.Д. Цвиркуна* «Управление развитием: структурный анализ, задачи, устойчивость» посвящен вопросам формирования методологии управления развитием. Работа посвящена фундаментальным проблемам теории управления в системах, мотивированных интересами. Выполняется структурный анализ, по результатам которого формируется методология управления со структурой марковского процесса коллективных решений. Формулируются принципы и аксиоматика коллективных решений, условия разрешимости компромиссом, постановки задач, условия устойчивости процесса управляемого развития. Развита методология управляемого развития, применимая к любым системам с объектом интересов, имеющим материальный носитель. Из ее методологических и конструктивных построений вытекает набор итоговых результатов и выводов, имеющих фундаментальный характер.

В докладе чл.-корр. РАН *А.Ф. Резчикова*, д-ра техн. наук *В.А. Кушникова* и д-ра техн. наук *В.А. Твердохлебова* «Модели и методы управления развитием крупномасштабной системы», представленном *В.А. Твердохлебовым*, управление развитием крупномасштабной системы рассматривается как управление изменениями правил управления событиями, составляющими процесс функционирования системы. Разработаны классификации процессов, процедур изменений правил управления и на основе совмещения классификаций построена классификация методов управления развитием крупномасштабных систем. Разработаны структуры управления развитием сложных систем.

В докладе д-ра техн. наук *О.С. Сухарева* «Влияние структуры крупномасштабной системы на ее рост: проблема запуска роста в России» представлен еще один аспект управления развитием на примере социально-экономической системы. Рассматривается проблема структурной динамики экономической системы, определения влияния ее

элементов на темп экономического роста. Проведен структурный анализ динамики компонентов ВВП России, который показал, что важно для верификации политики экономического роста на современном этапе развития. Применение методологии структурного анализа позволяет устанавливать связи в рамках экономической динамики и определять факторы, способные эту динамику изменить. По существу, речь идет о создании макроуправления крупной системой по ее структурным элементам, образующим общий вектор движения экономики.

Доклад чл.-корр. РАН *В.И. Суслова* «Пространственный аспект стратегии социально-экономического развития России» посвящен разработке и анализу перспективных и сверхперспективных прогнозов развития страны. Российское экономическое пространство крайне неоднородно, фрагментировано, моноцентрично. При определении позиций отдельных макрорегионов в экономике страны, общего качества экономического пространства пространственный фактор выходит на первые роли. Создана экспериментальная расчетная модель, позволяющая сконструировать несколько сценариев пространственного развития на средне- и долгосрочную перспективу. Разработаны такие сценарии, из которых можно сформировать целевой сценарий комплексного характера, учитывающий определенное сочетание опорных сценариев некоторых типов, например, межрегиональные экономические отношения, предполагающие сценарий, в котором достигается оптимум критерия «справедливости» пространственных взаимоотношений.

В докладе д-ра экон. наук *В.Н. Лившица*, канд. экон. наук *Т.И. Тищенко* и канд. экон. наук *М.П. Фроловой* «Крупномасштабные социально-экономические катастрофы: причины и последствия» в качестве примера катастрофического социально-экономического кризиса представлены события 1991 г. в России, когда системы централизованного планирования и управления в производственной и социальной сферах были заменены «саморегулирующимся рыночным фантомом». Показаны основные причины кризиса: неправильный выбор при реформировании российской экономики и социума базовой макроэкономической теории, когда была принята неправомерная и неэффективная смена «советской коллективистской философии жизни и идеологии хозяйствования» на индивидуалистическую монетаристско-псевдорыночную; господство в управлении мифов типа «частное управление всегда эффективнее государственного» и т. п. Путь восстановления — в



возвращении нормальных ценностей социального общества.

Доклады чл.-корр. РАН *Н.А. Махутова*, д-ра техн. наук *А.Ф. Бермана*, д-ра техн. наук *О.А. Николайчука* «Алгоритм самоорганизующегося процесса обоснования свойств технического состояния сложных опасных объектов», д-ра техн. наук *А.Г. Полетыкина*, канд. техн. наук *Е.Ф. Жарко*, *Н.Э. Менгазетдинова*, канд. физ.-мат. наук *В.Г. Промыслова* «Новое поколение систем верхнего уровня и концепция Industry 4.0», д-ра техн. наук *В.В. Кульбы* «Сценарный анализ в управлении геополитическим комплексным информационным противоборством», д-ра техн. наук *В.Н. Буркова*, д-ра техн. наук *И.В. Бурковой*, *В.Б. Барковской* «Управление рисками проектов на основе качественных оценок» были посвящены различным аспектам безопасности разных систем и объектов, разным методикам и моделям исследования проблем управления рисками.

Так, в интересном выступлении *В.Н. Буркова* управление рисками рассматривается как одна из важных функциональных областей в методологии управления проектами. Предлагаются элементы теории управления рисками на основе качественных оценок, дается определение качественных оценок сложных рисков, формулируются задачи управления рисками путем их снижения, передачи или уклонения от риска, создания резервов и др. Во всех случаях реагирования возникает задача снижения уровня риска с минимальными затратами и потерями. Дается определение основных качественных характеристик рисков и сложных рисков, рассматривается ряд задач управления рисками и предлагаются методы решения. Приняв значения качественных оценок рисков, руководитель проекта получает строгий математический инструментарий для принятия решений по снижению рисков.

В докладе д-ра техн. наук *В.П. Кутахова*, канд. техн. наук *С.И. Пляскоты* «Информационное взаимодействие в крупномасштабных робототехнических авиационных системах» тема рисков и безопасного функционирования сложных технических систем развивается на материале исследования информационного взаимодействия в группах беспилотных летательных аппаратов — подвижных робототехнических элементов крупномасштабных авиационных робототехнических систем в существенно переменных условиях применения. Рассматриваются вопросы построения системы управления в группе, организации информационного обмена между робототехническими элементами группы. Сформулирован перечень задач при формировании группового поведения элементов как

задачи управления поведением группы, ее реконфигурации и формирования структуры информационного обмена. Выделена и подробно исследована задача целераспределения и взаимного целеуказания между элементами группы. Формализована задача оптимизации взаимного целеуказания в пространственно-распределенной группе разнородных информационных систем.

В докладе группы ведущих ученых РАН д-ра техн. наук *А.А. Зацаринного*, д-ра техн. наук *К.К. Колина*, д-ра техн. наук *Н.И. Ильина*, д-ра психолог. наук *В.Е. Лепского*, д-ра физ.-мат. наук *Г.Г. Малинецкого*, д-ра техн. наук *А.Н. Райкова*, канд. физ.-мат. наук *Б.Б. Славина* «Система распределенных ситуационных центров развития», представленном *А.Н. Райковым*, систематизируется философское, психологическое и технологическое обеспечение задачи стратегического развития в новой парадигме научных исследований — переходе к наукам с интенсивным использованием данных. Создается национальная исследовательская инфраструктура, для повышения эффективности принимаемых управленческих решений на всех уровнях государственного управления создается система распределенных ситуационных центров. Успешное решение этих задач будет способствовать обеспечению более высокого уровня национальной безопасности и качества жизни.

Проблеме информационного обеспечения различных аспектов «модели интересов» руководства корпорации, в том числе упорядочению информационных потоков и интегрированной обработке данных различного типа, посвящен доклад *М.С. Костина*, канд. техн. наук *В.М. Розина*, *М.Ю. Стерхова*, *А.А. Царева* «Управленческие системы — сборка вместо разборки». Предлагается технология ASID (assembling instead of development), в которой для автоматизированного проектирования управленческих систем независимо от сферы и масштабов деятельности используется набор готовых универсальных решений. В основе технологии — принцип «Сборка вместо разработки», реализуемый с помощью интерактивной матрицы запросов, которая обеспечивает использование и адаптацию апробированных на практике готовых решений при автоматизированной подготовке прототипа разрабатываемой системы.

С помощью разработанной технологии автоматизированного проектирования управленческих систем и быстрой адаптации готовых функциональных модулей — визуально-интерактивных панелей, обеспечивающих многоаспектный анализ ситуации, независимо от сферы и масштабов деятельности, руководство корпорации получает возможность оперативно улавливать опасные сим-

птомы и быстро, комплексно визуально оценивать ситуацию «в целом». Это становится одним из важнейших направлений при развертывании управленческих систем.

Секционная работа конференции MLSD'2017 была проведена активно, особенно в секциях управления развитием крупномасштабных систем, инвестиционной, имитационно-оптимизационной, транспортной, аэрокосмической и мониторинговой. Отмечено участие большого числа молодых ученых и специалистов — исследователей, аспирантов и студентов старших курсов вузов и содержательное общение аудитории с докладчиками.

Участниками круглых столов, проведенных под председательством руководителей направлений, были отмечены широта тематики конференции и возрастающая потребность в научной координации работ, проводимых по проблематике ежегодной международной конференции «Управление развитием крупномасштабных систем».

Конференция прошла под знаком разработки и реализации Стратегии научно-технологического развития. Большое внимание уделялось вопросам долгосрочного планирования, развития системы научно-технологического прогнозирования, оценки качества крупномасштабных проектов развития регионов, отраслей народного хозяйства, повышения качества экспертизы для принятия эффективных решений в области экономического развития

страны, развития организационных форм научной, научно-технической и инновационной деятельности, государственного управления.

На заключительном заседании был представлен и обсужден проект решения конференции и принят ряд предложений.

По результатам десятой международной конференции «Управление развитием крупномасштабных систем» MLSD'2017 опубликованы Материалы, включающие в себя пленарные доклады и тезисы секционных докладов (в 2-х т.), и Сборник трудов, индексированный в РИНЦ (в 2-х т.), 134 доклада зарегистрированы в IEEE (индексация в системе Scopus). Кроме того, наиболее интересные в смысле научных результатов доклады рекомендованы для публикации в журналах «Автоматика и телемеханика» (тематический выпуск), «Проблемы управления», «Системы управления и информационные технологии» и др.

*Председатель Оргкомитета А.Д. Цвиркун,
член Оргкомитета Э.Г. Прохорова*

Цвиркун Анатолий Данилович — д-р техн. наук, зав. отделом,
✉ tsvirkun@ipu.ru,

Прохорова Элла Григорьевна — науч. сотрудник,
✉ proipu@yandex.ru,

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва.



advances in
systems science
and applications

Содержание журнала *Advances in Systems Science and Applications,* **Vol. 17, No 4 (2017)**

Predictive Feedback Control Method for Stabilization of Continuous Time Systems

Lina Ahmed Shalby

P. 1—13

Selective Attention Compensates Detrimental Effect of Auditory Noise on Performance in Posner Experiment

Borzou Alipourfard, Kamyar Arbabifard, Fatemeh Bakouie, Somayeh Sadat Hashemi Kamangar, Shahriar Gharibzadeh

P. 14—21

Hybrid PAPR Reduction Scheme for Universal Filter Multi-Carrier Modulation in Next Generation Wireless Systems

Pooja Rani, Himanshu Monga, Silki Baghla

P. 22—33

The Development of the agent-based demography and migration model of Eurasia and its supercomputer implementation

Albert R. Bakhtizin, Valery Makarov, Elena Sushko, Gennady Sushko

P. 34—45

Threshold analysis of a stochastic epidemic model with delay and temporary immunity

Rui Xue, Fengying Wei

P. 46—60

Price of Anarchy for Maximizing the Minimum Machine Load

Yulia Vasil'evna Chirkova

P. 61—77

Developing a Strategy of Environmental Management for Electric Generating Companies Using DEA-Methodology

Svetlana Valerievna Ratner, Pavel Dmitrievich Ratner

P. 78—92

Тексты статей в свободном доступе на сайте <http://ijassa.ipu.ru/ojs/ijassa>

Александр Александрович ДОРОФЕЮК 1941—2018

18 февраля 2018 г. после тяжелой и продолжительной болезни ушел из жизни выдающийся российский и советский ученый, организатор науки, член редколлегии журнала «Проблемы управления» Александр Александрович Дорофеюк.

Александр Александрович пришел в Институт автоматики и телемеханики АН СССР (так тогда назывался Институт проблем управления имени В.А. Трапезникова РАН) в 1962 г., будучи студентом МФТИ, а в 1968 г. был принят на работу по окончании аспирантуры МФТИ в лабораторию Марка Ароновича Айзермана.

В 1969 г. А.А. Дорофеюк защитил кандидатскую диссертацию «Методы распознавания образов без учителя и их применение» (научный руководитель — М.А. Айзерман). В последующие годы под руководством Э.М. Бравермана он продолжал исследования в области интеллектуальных методов анализа сложно организованных данных, разработки новых алгоритмов интеллектуального анализа информации, их теоретического исследования и применения для решения широкого спектра прикладных задач. Когда в 1977 г. Э.М. Бравермана не стало, Александр Александрович возглавил это направление исследований.

В последующее десятилетие исследование задач классификационного анализа было связано с вариационным подходом. Начиная с 1980—1981 гг., вариационный подход к исследованию алгоритмов классификационного анализа применялся для случая размытой классификации. К концу 1980-х гг. группой А.А. Дорофеюка был получен ряд важных теоретических результатов, позволивших свести к задаче классификационного анализа целый ряд других задач интеллектуального анализа информации.

Одновременно велись интенсивные работы по применению разработанных алгоритмов для решения ряда прикладных задач. Наиболее масштабными и эффективными оказались проекты «Зап-



часть», «Автоперевозки», «Региональное здравоохранение» и ряд других. По результатам выполненных исследований в 1981 г. А.А. Дорофеюк защитил докторскую диссертацию «Методы автоматической классификации и проблема исследования организационных систем управления». В 1984 г. ему было присвоено ученое звание профессора.

В 1988 г. в Институте создается новая лаборатория, заведующим которой назначается А.А. Дорофеюк. Сотрудниками лаборатории под его руководством разработаны методы перекрестной экспертизы, экспертно-классификационного анализа данных и структурного прогнозирования.

Последние годы жизни А.А. Дорофеюк руководил работами по целевому гранту РФФИ и принимал участие в работах по ряду других грантов РФФИ, связанных с совершенствованием системы управления Национальным медицинским исследовательским центром нейрохирургии имени академика Н.Н. Бурденко.

Долгие годы А.А. Дорофеюк был членом Ученого совета ИПУ РАН и председателем одной из его секций.

Александр Александрович был человеком умным, небезразличным и эмоциональным. От него могло и достаться, порою, несправедливо. Но одновременно он был человеком широким и участливым, умел помогать людям как никто другой. К нему просто потоком шли те, у кого были почти неразрешимые проблемы — не только в науке, но и по жизни. И он помогал всем. В науке, предлагая новые идеи. Решая непробиваемые административные вопросы. Подыскивая самых лучших врачей. И многое, многое другое. Память об Александре Александровиче Дорофеюке сохранится у всех, кто соприкасался с ним.

*Сотрудники Института проблем управления
им. В.А. Трапезникова РАН
Редколлегия
Редакция*



Институт проблем управления им. В.А. Трапезникова РАН осуществляет подготовку (по очной и заочной форме) научных кадров в аспирантуре

Направление подготовки: 01.06.01 — «Математика и механика»

Специальность:

- 01.01.02 — Дифференциальные уравнения, динамические системы и оптимальное управление.

Направление подготовки: 09.06.01 — «Информатика и вычислительная техника»

Специальности:

- 05.13.01 — Системный анализ, управление и обработка информации (по отраслям) — по техническим наукам;
- 05.13.01 — Системный анализ, управление и обработка информации (в отраслях информатики, вычислительной техники и автоматизации) — по физико-математическим наукам;
- 05.13.05 — Элементы и устройства вычислительной техники и систем управления — по техническим наукам;
- 05.13.06 — Автоматизация и управление технологическими процессами и производствами (по отраслям) — по техническим наукам;
- 05.13.10 — Управление в социальных и экономических системах — по техническим наукам;
- 05.13.11 — Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей — по техническим наукам;
- 05.13.12 — Системы автоматизации проектирования (по отраслям) — по техническим наукам;
- 05.13.15 — Вычислительные машины, комплексы и компьютерные сети;
- 05.13.18 — Математическое моделирование, численные методы и комплексы программ — по техническим наукам.

Направление подготовки: 38.06.01 — «Экономика»

Специальности:

- 08.00.05 — Экономика и управление народным хозяйством (по отраслям и сферам деятельности, в том числе управление инновациями);
- 08.00.13 — Математические и инструментальные методы экономики.

Сроки обучения. В очной форме аспирантуры — 4 года (по направлению подготовки 38.06.01 — 3 года);
В заочной форме аспирантуры — 5 лет (по направлению подготовки 38.06.01 — 4 года).

Прием в аспирантуру. Все поступающие в аспирантуру проходят предварительное собеседование с предполагаемым научным руководителем. Результаты собеседования учитываются при решении вопроса о допуске поступающего к вступительным экзаменам.

Прием вступительных испытаний в аспирантуру проводится в конце августа — сентябре.

Прием заявлений и документов в аспирантуру — с 25 июня по 1 августа. Зачисление в аспирантуру до 1 октября. Документы и заявления подаются в Отдел докторантуры и аспирантуры ИПУ РАН. Вступительные экзамены по специальной дисциплине и иностранному языку проводятся до середины сентября. Вступительные экзамены могут быть проведены на русском и иностранных языках.

В Институте **читаются лекции для аспирантов** по следующим программам:

- «Эконометрический анализ экономического развития» — д-р экон. наук, профессор Нижегородцев Р.М.;
- «Теория автоматического управления» — д-р техн. наук, профессор С.А. Краснова;
- «Теория робастного управления» — д-р техн. наук, профессор А.П. Курдюков;
- «Теория многоагентных систем» — д-р техн. наук, профессор Р.П. Агаев;
- «Идентификация систем управления» — канд. техн. наук Макаров В.В.;
- «Методы дискретной оптимизации и управление проектами» — д-р физ.-мат. наук, профессор РАН М.В. Губко;
- «Управление и искусственный интеллект» — канд. техн. наук А.В. Макаренко;
- «Технологические основы управления производством продукции на промышленных предприятиях» — д-р техн. наук, профессор А.Л. Генкин;
- «Вычислительные машины, комплексы и компьютерные сети» — д-р техн. наук, профессор В.М. Вишневский;
- «Системы автоматизации проектирования» — д-р техн. наук, профессор А.В. Толок.

МЕРОПРИЯТИЯ IFAC — Международной федерации по автоматическому управлению

Мероприятие	Дата	Место проведения
IFAC Conference on Technology, Culture and International Stability — 18th TECIS 2018	13—15 Sep 2018	Baku, AZERBAIJAN
INSTICC, IFAC et al. International Joint Conference on Computational Intelligence — 10th IJCCI 2018	18—20 Sep 2018	Seville, SPAIN
IFAC Conference on Engine and Powertrain Control, Simulation and Modeling — 5th E-COSM 2018	20—22 Sep 2018	Changchun, CHINA
IFAC Workshop on Control Applications of Optimization — 17th CAO 2018	15—19 Oct 2018	Yekaterinburg, RUSSIA
Conference on Latin American Conference on Automatic Control (in cooperation with IFAC) — XVIII CLCA 2018	24—26 Oct 2018	Quito, ECUADOR
OTM/IFAC/IFIP et al. International Workshop on Enterprise Integration, Interoperability and Networking — 13th EI2N 2018	24—25 Oct 2018	Valletta, MALTA
IFAC Conference on Analysis and Control of Chaotic Systems — 5th CHAOS 2018	30 Oct —1 Nov 2018	Eindhoven, NETHERLANDS
IFAC Conference on Cyber-Physical and Human Systems — 2nd CPHS 2018	13—15 Dec 2018	Miami (FL), USA
IFAC Symposium on Dynamics and Control of Process Systems, including Biosystems — 12th DYCOPS 2019	23—26 Apr 2019	Florianopolis, BRAZIL
IFAC Workshop on Control of Systems Governed by Partial Differential Equations — 3rd CPDE 2019	22—24 May 2019	Oaxaca, MEXICO
SICE, ISCIE, JSME, IEEE, IFAC Conference on Asian Control Conference (in cooperation with IFAC) — 12th ASCC 2019	9—12 Jun 2019	Kitakyushu, Fukuoka, JAPAN
IFAC Workshop on Control of Smart Grid and Renewable Energy Systems — CSGRES 2019	10—12 Jun 2019	Jeju, REPUBLIC OF KOREA
IFAC Symposium on Advances in Automotive Control — 9th AAC 2019	25—28 Jun 2019	Orlians, FRANCE
EUCA, IFAC, IEEE CSS Conference on European Control Conference (in cooperation with IFAC) — ECC 2019	9—12 Jun 2019	Naples, ITALY
IFAC Symposium on Intelligent Autonomous Vehicles — 10th IAV 2019	3—5 Jul 2019	Gdansk, POLAND
IFAC Workshop on Thermodynamic Foundations for a Mathematical Systems Theory — 3rd TFMST 2019	8—10 Jul 2019	Louvain-la-Neuve, BELGIUM
Conference on American Control Conference (in cooperation with IFAC) — ACC 2019	10—12 Jul 2019	Philadelphia, Pennsylvania, USA
IFAC, IEEE, IFIP, IFORS et al. Conference on Manufacturing Modelling, Management and Control — 9th MIM 2019	28—13 Aug 2019	Berlin, GERMANY
IFAC Symposium on Advances in Automotive Control — 9th AAC 2019	25—28 Jun 2019	Orlians, FRANCE

Более полная информация доступна на сайте www.ifac-control.org

CONTENTS & ABSTRACTS

UPPER ESTIMATES OF LARGE DEVIATIONS IN LINEAR SYSTEMS IN PRESENCE OF UNCERTAINTY 2

Kvinto Ya.I., Khlebnikov M.V.

The linear dynamic system is considered in the presence of uncertainty in its matrix. Upper estimates of deviations in linear dynamic systems are obtained on the basis of linear matrix inequalities technique, and the problem is investigated of deviations minimization in linear control systems by means of static linear state feedback. Numerical simulations demonstrated the low degree of conservatism of the estimates obtained.

Keywords: linear dynamic system, uncertainty, linear matrix inequalities, large deviations, Lyapunov function.

YOUTH GRANTS AS A MANAGEMENT TOOL FOR SCIENCE DEVELOPMENT: EFFECTIVENESS ANALYSIS 8

Klochkov V.V., Rozhdestvenskaya S.M.

The analysis is conducted of the impact of financial support in the form of grants for youth on the effectiveness of the selection of more able scientists. The influence is evaluated of the scientist financial security on a time he can devote to science, on his performance during the period of grant support and, ultimately, on the quality of the scientists' selection on the basis of their ranking. The requirements for the level of grant support are determined. The impact of the duration of the scientists' support and performance monitoring period on the selection quality is assessed. The desirable time and cost parameters of young scientists grant support are determined.

Keywords: development management, grants, young scientists, life cycle, productivity, ranking, efficiency, errors of the first and second kind.

COST-REDUCING FUNDING MECHANISMS 17

Shchepkin A.V.

The task is considered of determining the amount of financing for several jobs constituting the total production order. The contractor of each job is a monopolist in a respective field and has no competitors. Tasks financing funds are given. Mechanisms are investigated providing the true information acquisition on the cost of each job fulfillment.

Keywords: costs, profit, profitability, cash advance.

CONFLICT-FREE SELF-ROUTING FOR THREE-DIMENSIONAL GENERALIZED HYPERCUBE . . . 26

Podlazarov V.S.

The system-area network is proposed in the form of nonblocking three-dimensional generalized hypercube. The new hypercube structure is designed and the algorithm is suggested for it of conflict-free making of direct paths between the nodes by the means of self-routing in them.

Keywords: system-area network, generalized hypercube, complete multirouting, switching features, nonblocking network, conflict-free self-routing, direct paths.

METHOD OF CONTROLLING TEST CONSTRUCTION FOR DIGITAL SYSTEMS WITH MINIMAL QUASICOMPLETE GRAPH STRUCTURE (ON EXAMPLE OF GRAPH BY DIMENSION 11×11) 33

Vedeshenkov V.A.

The new method is worked out of controlling test construction for digital systems with minimal quasicomplete graph structure. Special tests were formed for n fragments of digital system, each consisting of a commutator and $2m$ communication lines. Later the special tests check-ups were divided in n groups with general testing abonents. The set of nm check-ups represents the full minimal test for single defective components of digital systems of the structure reviewed. The example is presented of controlling test construction for digital system by dimension 11×11 . It is noted that the method suggested can be used for the analysis of digital systems of the structure reviewed with other parameters of graph.

Keywords: digital system, fragment, abonent, commutator, communication line, minimal quasicomplete graph, controlling test, special test.

CREATION OF INFORMATION-TECHNOLOGICAL RESERVE IN DISTRIBUTED DATA PROCESSING SYSTEMS 40

Somov S.K.

The main stages are presented of creating an information-technological data reserve in distributed automated informational control systems. The reserve is formed on the basis of the results of the analysis of user domains, regular requests to the system and the analysis of the execution sequence of queries processing procedures. The tasks are set of the reserve structure design and of the optimal allocation of reserve copies among computer network nodes. It is noted that this type of reserve allows reducing the processing time of typical, regular requests to the system due to the fact that the reserve contains the data sets, prepared beforehand and later used at requests processing.

Keywords: distributed data processing system, computer network, information-technological reserve.

MANAGING THE CYBERSECURITY ARCHITECTURE OF AUTOMATED PROCESS CONTROL SYSTEMS OF NUCLEAR POWER PLANTS 47

Babaev D.I., Poletikin A.G., Promyslov V.G., Timofeev M.Yu.

Considered are the methods of risk assessment, cybersecurity levels assignment, the requirements to protection measures and to cybersecurity architecture design. The main tasks are formulated of ensuring the cybersecurity. The principles are outlined of constructing the security architecture. The concrete example of the practical implementation of the security architecture is discussed.

Keywords: information, cybersecurity, security, automated control systems, nuclear power plant.

METHOD OF RECURRENT AND Z-RECURRENT CONTROL OF COMPLEX SYSTEM FUNCTIONING 56

Rezchikov A.F., Tverdokhlebov V.A.

The method is developed of constructing the rules of controlling the interactions between complex system functioning processes. Processes are considered as sequences of cause-and-effect related events, specified using their properties indicators values. The structure of control rules is presented on the basis of the recurrent and first introduced Z-recurrent definitions of sequences in the form of functional dependencies of the event properties indicators values. It is noted that the models developed of the events, processes and complex system functioning are applicable on the whole when solving the problems of control, monitoring, and diagnosis.

Keywords: functioning of complex system, process, event, control rule, recurrent and Z-recurrent definition of sequence.

LOCOMOTIVE ASSIGNMENT GRAPH MODEL FOR FREIGHT TRAFFIC ON LINEAR SECTION OF RAILWAY. THE PROBLEM OF FINDING A MAXIMAL INDEPENDENT SCHEDULE COVERAGE 65

Zhilyakova L.Yu., Kuznetsov N.A., Matiukhin V.G., et al.

The paper is devoted to the formal statement and solution of a problem arising when assigning the locomotives for freight transportation realization in accordance with preset schedule. The goal is to determine whether the number of locomotives is sufficient at a specified initial allocation of them to perform all transport operations. The solution is presented in the form of an algorithm that builds the coverage of the schedule: the complete one, if it exists, or else the partial one being the maximal independent. The theorem is proved on one-to-one correspondence between the existence of the complete coverage and the sufficiency of the number of locomotives.

Keywords: graph model, network flows, locomotive assignment, freight rail transportation.

THE 10th INTERNATIONAL CONFERENCE «MANAGEMENT OF LARGE-SCALE SYSTEM DEVELOPMENT MLSD'2017». 76

Aleksander Dorofeyuk 81

СИСТЕМА ОПЕРАТОР



Продукты и технологии

- SCADA (с адаптацией под заказчика);
- операционные системы для АСУ (поставка готовых и изготовление по ТЗ);
- технологии интеграции АСУ из разнородных систем автоматики;
- технологии и продукты для разработки и контроля безопасных и защищенных АСУ ТП;
- простые и сложные пульты управления для АСУ с 1 млн. сигналов и более;
- средства оценки кибербезопасности;
- поддержка работы на процессорах X86 и ARM, RISK;
- разработаны в соответствии с POSIX и ЕСПД.

Особенности

- собственная разработка;
- "атомное" качество;
- масштабируемость (1 млн. сигналов и более);
- долговременная поддержка российских разработчиков;
- доступные исходные коды;
- кибербезопасность;
- поставки за рубеж.

Услуги

- разработка программно-технических компонент и комплексов: верхний/средний уровень / встраиваемые системы;
- верификация и валидация АСУ, включая исходные тексты, по качеству/информационной безопасности;
- разработка ПО для проектного анализа контролепригодности, технической эффективности, надежности, техногенной безопасности технических систем;
- анализ эксплуатационной готовности технологических комплексов;
- анализ кибербезопасности;
- консультации специалистов;
- обучение.

Разработки внедрены на АЭС "Бушер", Исламская Республика Иран и АЭС "Куданкулам", Республика Индия.

Зарегистрировано в ЕДИНОМ РЕЕСТРЕ российских программ для электронных вычислительных машин и баз данных, №3290 от 29 марта 2017 года

ИНСТИТУТ ПРОБЛЕМ УПРАВЛЕНИЯ им. В.А. ТРАПЕЗНИКОВА РОССИЙСКОЙ АКАДЕМИИ НАУК
ул. Профсоюзная, 65, 117997, Москва, Россия.

телефон: +7 (495) 334-89-10; факс: +7 (495) 334-93-40, +7 (499) 234-64-26

e-mail: ipu31@mail.ru, poletik@inbox.ru

<http://www.ipu.ru>

