



**CONTROL SCIENCES**

**Научно-технический  
журнал**

6 номеров в год

ISSN 1819-3161

**УЧРЕДИТЕЛЬ**

Учреждение Российской  
академии наук

Институт проблем управления  
им. В.А. Трапезникова РАН

**Главный редактор**

**Д.А. Новиков**

**Заместители главного  
редактора**

**Л.П. Боровских, Ф.Ф. Пащенко**

**Редактор**

**Т.А. Гладкова**

**Выпускающий редактор**

**Л.В. Петракова**

**Издатель**

**ООО «СенСидат-Контрол»**

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  
ООО «Авансед Солюшнз»

Отпечатано в ИПУ РАН

Заказ № 96

Подписано в печать  
21.07.2010 г.

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

Свидетельство о регистрации  
ПИ №77-11963 от 06 марта 2002 г.

Журнал входит в Перечень ведущих  
рецензируемых журналов и изданий,  
в которых должны быть опубликованы  
основные научные результаты  
диссертаций на соискание ученой  
степени доктора и кандидата наук

Подписные индексы:  
**80508** и **81708** в каталоге Роспечати  
**38006** в объединенном каталоге  
«Пресса России»

Цена свободная

© Учреждение Российской  
академии наук  
Институт проблем управления  
им. В.А. Трапезникова РАН

# ПРОБЛЕМЫ УПРАВЛЕНИЯ

## 4.2010

### СОДЕРЖАНИЕ

#### Обзоры

Дубина И.Н. Зарубежные исследования в области теоретико-игрового анализа инноваций . . . . . 2

#### Математические проблемы управления

Зак Ю.А. Решение обобщенной задачи Джонсона с ограничениями на сроки выполнения заданий и времена работы машин. Ч. 2. Приближенные методы решения . . . . . 12

#### Управление в социально-экономических системах

Абрамов А.П. Инвестиции и сбалансированный рост в модели децентрализованной экономики. . . . . 20

Павлов О.В. Принятие инвестиционных решений на основе теории оптимального управления дискретными системами . . . . . 27

Дорофеев Ю.А., Мандель А.С. Структурно-классификационная методология оценки эффективности функционирования жилищно-коммунального хозяйства крупного города (на примере Москвы) . . . . . 34

Куливец С.Г. Моделирование конфликтных ситуаций с несогласованными представлениями у агентов на основе игр на линейных когнитивных картах. . . . 42

#### Информационные технологии в управлении

Губко М.В., Даниленко А.И. Математическая модель оптимизации структуры иерархического меню . . . . . 49

Асратян Р.Э. Взаимодействие клиента и сервера в Интернет-службе обработки удаленных пакетов файлов . . . . . 59

#### Управление техническими системами и технологическими процессами

Попова О.М., Усов И.Ю. Оптимизация развития системообразующей электрической сети с помощью геоинформационных технологий . . . . . 66

Ефимов А.Е., Опалев О.Л., Ядыкин И.Б. Определение коэффициентов влияния параметров режима электроэнергетической системы на ее устойчивость. . . . . 74

Финягина В.И. Расчет аппроксимирующих функций двумерных температурных полей в задачах управления подвижными источниками воздействий . . . . . 79

\* \* \*

Contents and abstracts. . . . . 86

# ЗАРУБЕЖНЫЕ ИССЛЕДОВАНИЯ В ОБЛАСТИ ТЕОРЕТИКО-ИГРОВОГО АНАЛИЗА ИННОВАЦИЙ<sup>1</sup>

И.Н. Дубина

Представлен обзор современных исследований, посвященных применению методов теории игр для анализа различных аспектов инновационной деятельности. Охарактеризованы модели поведения инновационной фирмы в конкурентной среде, стратегии передачи прав на объекты интеллектуальной собственности, варианты организационно-экономических отношений участников инновационного консорциума и вопросы формирования инновационной политики.

**Ключевые слова:** теория игр, инновация, конкуренция, лицензирование, инновационный консорциум.

## ВВЕДЕНИЕ

Принципы и методы теории игр нахождения оптимальных решений в ситуациях конкуренции, неопределенности и риска широко применяются сегодня в самых разных областях, в том числе для моделирования и анализа инновационной активности фирм. В данной статье представлен обзор работ, посвященных применению методов теории игр для анализа различных аспектов инновационной деятельности и опубликованных в зарубежных изданиях в течение последних 30 лет. Представленное исследование основано на использовании полнотекстовых баз данных, содержащих монографии, доклады, диссертации и статьи из нескольких тысяч академических журналов. В частности, использовались такие ресурсы, как EBSCO, Cambridge Scientific Abstract (CSA), Electronic Information Service, Web of Science (Science Citation Index and Social Sciences Citation Index), ScienceDirect, SpringerLink, Emerald Management Xtra, JSTOR (Journal Storage), LexisNexis Academic, NetLibrary, Oxford Scholarship Online, Web of Knowledge и др.

<sup>1</sup> Работа подготовлена при финансовой поддержке Российского фонда фундаментальных исследований (РФФИ), грант № 10-06-98008-р\_Сибирь\_а.

Доступ к информационным ресурсам, использованным для проведения данного исследования, был предоставлен Центрально-Европейским университетом (г. Будапешт) по гранту CEU Special Program Office, Special and Extension Programs.

Обзоры статей, посвященных теоретико-игровому моделированию инновационной активности и опубликованных в 1970—1980-х гг., представлены в работах [1, 2], которые показывают, что за этот период было опубликовано всего около 30 статей, непосредственно посвященных анализу инноваций с использованием теоретико-игровых моделей. За последние 10 лет число публикаций по этой теме значительно увеличилось. В частности, в базе EBSCO, содержащей тексты статей из более, чем 10 тыс. периодических изданий и почти 11 тыс. монографий и диссертаций, зарегистрировано всего 6 работ, связанных с инновациями и теорией игр и опубликованных в 1960-х гг., 20 — в 1970-х, 10 — в 1980-х, 29 — в 1990-х и 105 — в 2000-х гг.

Однако обращает на себя внимание тот факт, что «область пересечения» теории игр и исследований в области инноваций остается очень узкой. Например, в базе EBSCO зарегистрировано всего 170 статей, в тексте которых встречаются слова, релевантные терминам «теория игр» (game theory) и «инновации» (innovation), и опубликованных в период с 1962 по 2009 г. Но в этой же базе зарегистрировано 25 056 статей со словами «теория игр» и 807 989 статей со словами, близкими к слову «инновации». Чтобы исключить статьи со «случайным» попаданием в текст популярного слова «инновация», отбирались работы, в которых слова «теория игр» и «инновации», а также близкие к ним по контексту (например, R&D), включены либо в заголовок статьи, либо в раздел ключевых



слов, либо в аннотацию. Таких статей в EBSCO зарегистрировано всего 55. В этой же базе зарегистрировано 21 850 статей, непосредственно относящихся к теории игр, 665 581 — к различным аспектам инновационной деятельности и 82 191 — к анализу НИОКР (R&D). Такой результат с определенной степенью условности может быть интерпретирован следующим образом: лишь в около 0,25 % статей по теории игр прямо рассматриваются возможности ее применения к инновационной деятельности и менее чем в 0,01 % статей, связанных с обсуждением инноваций, применяются методология и инструментарий теории игр. Сходные соотношения обнаруживаются и по другим исследовательским базам. Например, в Web of Knowledge зарегистрировано 1028 статей по теории игр, 9612 — по инновациям и всего 17 статей по тематике теории игр и инноваций. «Узость» пересечения двух исследовательских направлений представляется неожиданным результатом, принимая во внимание чрезвычайную популярность и широту исследований в области инноваций, а также тот факт, что теория игр уже давно рассматривается в качестве неотъемлемого методологического элемента серьезного экономического анализа.

Тем не менее, общее число англоязычных публикаций по рассматриваемой проблематике, зарегистрированных в ведущих исследовательских базах, составляет около 300. Эти работы представляют широкое разнообразие теоретико-игровых моделей разной сложности и прикладной направленности. На основе анализа этих публикаций в данной статье выделяются и иллюстрируются конкретными моделями основные направления приложений теории игр к анализу инноваций.

## 1. ОБЩАЯ ХАРАКТЕРИСТИКА ТЕОРЕТИКО-ИГРОВЫХ МОДЕЛЕЙ ИННОВАЦИЙ

Направления теоретико-игрового анализа инноваций, представленные в зарубежной литературе, концентрируются преимущественно на двух уровнях анализа «инновационных игр»:

- межорганизационные игры, в которых основными игроками выступают фирма, ее конкуренты, партнеры (например, инвесторы, дистрибьюторы, поставщики, исследовательские центры), клиенты, владельцы прав на объекты интеллектуальной собственности;
- мета-организационные игры, в которых игроками выступают государственный институт, определяющий инновационную политику, и агрегированный инновационный предприниматель.

На основе моделирования инновационных игр на межорганизационном уровне разрабатываются

оптимальные стратегии конкуренции и сотрудничества на высокотехнологических рынках, определяется оптимальный размер инвестиций в исследования и разработки (R&D), выбираются эффективные механизмы распределения результатов, получаемых от инновационных проектов, рассчитывается время вывода нового продукта на рынок, согласовываются параметры контрактов по передаче прав на использование новых технологий, «ноу-хау» и др. На мета-организационном уровне анализа инноваций с помощью теории игр предлагаются модели регулирования прав на интеллектуальную собственность, стимулирования инновационной активности в странах и регионах, развития инновационных кластеров и т. д.

В то же время, в зарубежной литературе почти не представлены теоретико-игровые исследования инновационной деятельности на внутрифирменном уровне, где в качестве основных игроков выступает менеджер проектов или администратор ресурсов и сотрудники фирмы — инициаторы и исполнители инновационных проектов. Основная задача теоретико-игрового моделирования на этом уровне заключается в разработке и обосновании оптимальных вариантов организационного и экономического взаимодействия участников внутрифирменной инновационной активности для эффективного стимулирования и активизации инновационной деятельности сотрудников, распределения полномочий и ресурсов при выполнении проекта, эффективной командной работы и т. д. [3]. Отметим, что в русскоязычной литературе теоретико-игровые модели и методы развития инноваций на внутрифирменном уровне представлены более обстоятельно [4].

Для моделирования инновационной деятельности применяются как кооперативные, так и некооперативные игры, причем последние представлены в литературе в значительно большей степени. Модели кооперативных игр используются преимущественно для решения проблем оптимального распределения результатов инновационных проектов среди их участников. Некооперативные игры с их фундаментальным принципом равновесия Нэша применяются для анализа очень широкого спектра проблем, включая инвестирование инновационных проектов, приобретение лицензий, выбор инновационных стратегий и др.

Во многих исследованиях специфика инновационных процессов «прячется» в используемой модели. Так, часто предполагается, что результатом инноваций становится снижение производственных затрат, повышение качества продукции и т. д., что в конечном счете приводит к получению дополнительной прибыли. Таким образом,

исследователи уходят от латентных и неизмеряемых или сложно-измеряемых переменных (например, уровня творческой активности сотрудников или интенсивности новаторской деятельности) к неким агрегированным параметрам (например, затратам на исследования и разработки, которые, как считается, отражают интенсивность новаторской деятельности). Снижение затрат, повышение качества, введение новых производственных и организационных решений, приносящих дополнительные эффекты, формализуется в виде «дельта»-параметров, включаемых в платежные функции игроков. Характерные особенности инновационных процессов часто моделируются вполне традиционно. Например, принимается, что прибыль фирмы, осуществляющей новый проект, составит  $H$  в случае успешного проекта с вероятностью  $x$  и  $L$ , если проект не будет удачным; тогда ожидаемый результат, включаемый в модель, вполне естественно определяется как математическое ожидание  $P = xH + (1 - x)L$  [5]. Довольно часто в модель вводятся случайные переменные (параметры-возмущения), характеризующие неопределенность спроса на новый продукт, непредвиденные изменения затрат на осуществление проекта и т. д.

Теоретико-игровой анализ инноваций в большинстве случаев основан на вполне традиционных моделях теории игр. Например, во многих теоретико-игровых моделях инноваций используется классическая игра Штакельберга, известная также как игра «лидер — последователь». Тем не менее, в ряд моделей их авторы включают параметры, характеризующие специфические особенности инновационной деятельности, например, «устаревание» знаний и технологий, накопление знаний, эффект влияния уровня новизны разрабатываемого продукта на затраты и спрос и др. В качестве примера приведем модель, в которой «уровень инновационности»  $\delta$  разрабатываемого продукта определяет затраты на его разработку  $c(\delta) = x\delta^2$ ,  $0 < x < 1$ , и предполагается, что новый продукт будет восприниматься потребителями как «лучший», что приведет к увеличению спроса на величину, пропорциональную  $\delta$  (желание потребителей платить больше за «новизну» зависит от «уровня новизны») [6]. Другой пример иллюстрирует отражение фактора риска в разрабатываемых моделях. В работе [7] функция  $F = 1 - e^{-ht}$  представляет вероятность того, что разработанный фирмой новый продукт сможет быть воспроизведен ее конкурентом через промежуток времени  $t$ . Если это произойдет, прибыль фирмы снизится с «монопольного уровня»  $P_m$  до уровня  $P_0$ , получаемого в условиях олигополистической конкуренции, следовательно,

величина  $[(1 - F)P_m + FP_0]$  может рассматриваться как ожидаемые денежные поступления фирмы-инноватора, а параметр  $h \geq 0$  — как показатель риска имитации (соответственно, предполагается, что значение этого параметра снижается при ужесточении законов, защищающих права интеллектуальной собственности).

## 2. МОДЕЛИРОВАНИЕ ПОВЕДЕНИЯ ИННОВАЦИОННЫХ ФИРМ В КОНКУРЕНТНОЙ СРЕДЕ

Большинство зарубежных публикаций, представляющих теоретико-игровые модели инновационной деятельности, ключевую роль отводят конкуренции. Классическая постановка проблемы инновационной конкуренции дана в работе [8]. Каждая из  $n$  конкурирующих фирм принимает решение об объеме ресурсов, инвестируемых в исследования и разработки (R&D), и направлении этих разработок. Инновационные стратегии могут различаться не только вероятностью их успеха в смысле коммерческой выгоды, но и тем, какую пользу можно извлечь из «провала» того или иного проекта. Выбор стратегии влияет не только на положение каждой фирмы на рынке, но и на технологическое развитие и производственную структуру рынка в целом. П. Дасгупта и Дж. Стиглиц [8, 9] предложили общую динамическую модель для анализа поставленной проблемы и взаимосвязи структуры рынка и технологических инноваций. Эта модель, несмотря на ее теоретичность и весьма абстрактный характер, послужила основой для последующей разработки широкого спектра моделей инновационной конкуренции.

Например, Р. Селлини и Л. Ламбертини [10] предложили динамическую модель для определения оптимальных стратегий по инвестированию инновационных разработок для увеличения степени дифференцированности выпускаемой продукции в условиях олигополистической конкуренции. Х. Смит и Л. Тригеоргис [11, 12] иллюстрируют конкурентные стратегии серией модельных примеров: от простых симметричных решений двух конкурентов по «запуску» новых проектов до кооперативных игр, моделирующих поведение участников инновационных альянсов. В работе [13] предложена теоретико-игровая модель с асимметричной информацией для анализа поведения двух конкурирующих фирм, разрабатывающих и стремящихся запатентовать новую технологию. Модель предполагает, что одна из фирм оценивает возможные преимущества получения патента выше, чем другая. Авторы показали, что вопреки «здравому смыслу», эта фирма будет инвестировать меньшие средства в развитие новой техноло-



гии и с меньшей вероятностью первой получит патент. Результаты моделирования проверялись в лабораторных экспериментах с различной мотивацией участников. Слабо мотивированные участники, действительно, «инвестировали» более активно и «выигрывали» чаще, подтверждая выводы авторов, полученные на основе моделирования. Другой пример моделирования инновационной конкуренции дан в статье [14]: фирмы из двух стран, производящие сходный продукт, принимают решение о режиме выхода на зарубежный рынок (прямые иностранные инвестиции, импорт, экспорт продукта и т. д.), размере инвестиций в собственные исследования и разработки, результатом которых станут технологии, сокращающие производственные затраты, и объеме продаж товара на каждом рынке (в каждой стране).

Игра  $n$  конкурирующих фирм, определяющих оптимальный уровень инновационной активности с учетом накопления знания, предложена в работе [15]. Если уровень инновационной активности  $i$ -й фирмы ( $i = 1, \dots, n$ ) составляет  $x_i$ , то фирма  $i$  получает доход, прямо пропорциональный  $x_i$  и обратно пропорциональный уровню суммарной активности всех ее конкурентов. Тогда прибыль, получаемая фирмой  $i$ , составит

$$\pi_i = B \frac{a_i x_i}{\sum_{j=1}^n a_j x_j} - C_i(\cdot),$$

где  $B$  — параметр, характеризующий потенциал рынка,  $C_i$  — затраты фирмы  $i$ , связанные с уровнем ее инновационной активности  $x_i$ .

При рассмотрении игры в динамике предполагается, что каждая фирма в период времени  $t$  выбирает уровень  $x_i$ , чтобы максимизировать ожидаемую прибыль в следующий период времени:

$$\pi_i^e(t+1) = B \frac{a_i x_i(t+1)}{a_i x_i(t+1) + \sum_{j \neq i} a_j x_j^e(t+1)} - C_i^{t+1}(\cdot), \quad (1)$$

где  $x_j^e(t+1)$  — уровень активности, ожидаемый фирмой  $i$  от конкурирующей фирмы  $j$  в период  $(t+1)$ .

Авторами принято, что затраты каждой компании линейно зависят от выбранного уровня инновационной активности  $x_i(t+1)$  и обратно про-

порциональны знанию, накопленному за период  $t$ ,  $z_i(t)$ :

$$C_i^{t+1}(x_i(t+1), z_i(t)) = \frac{x_i(t+1)}{1 + \lambda_i z_i(t)},$$

где параметры  $\lambda_i \geq 0$  характеризуют эффективность использования накопленных знаний и соответствующих технологий, уменьшающих затраты на разработку последующих инноваций.

Знания, получаемые фирмой  $i$  за период  $t$ , определяются как

$$X_i(t) = x_i(t) + \sum_{j \neq i} \vartheta_{ij} x_j(t),$$

где параметры  $\vartheta_{ij} \in [0, 1]$  отражают «утечку знаний» от фирмы  $j$  к фирме  $i$ . Если  $\vartheta_{ij} = 1$ , то знания, полученные фирмой  $j$  в результате ее инновационной активности, становятся полностью доступными фирме  $i$ . Другой предельный случай,  $\vartheta_{ij} = 0$ , соответствует невозможности переноса знаний и технологий от фирмы  $j$  к фирме  $i$  ( $\vartheta_{ij} = 0$  для всех  $i \neq j$  соответствует невозможности диффузии знаний).

Предполагается, что знания накапливаются с течением времени, т. е. предполагается, что фирма  $i$  извлекает пользу не только от знаний, полученных в период  $t$ , но и от знаний, полученных в прошлые периоды:

$$z_i(t) = x_i(t) + \sum_{j \neq i} \vartheta_{ij} x_j(t) + \rho z_i(t-1),$$

где параметр  $\rho$  характеризует скорость «устаревания» знаний. После идентификации всех параметров модели, для игры (1) может быть найдено численное решение.

Другая «классическая» проблема поведения инновационной фирмы в конкурентной среде заключается в выборе инновационной стратегии (например, осуществлять инновационную деятельность независимо или в сотрудничестве с другими фирмами, инвестировать средства в развитие новой технологии или осваивать технологию, представленной на рынке конкурентом и т. д.). Такой выбор зависит как от положения фирмы на рынке, так и от технологического развития рынка в целом [16]. Часто такая проблема рассматривается в варианте дилеммы «инновация или имитация». С одной стороны, фирма-имитатор следует за лидером-инноватором, поэтому не может получить краткосрочную монопольную прибыль, получаемую инноватором. С другой стороны, затраты имитатора на освоение новой технологии значительно меньше и составляют в среднем 65 % затрат

на разработку инновации [17]. Поэтому необходимо определить, какая стратегия (инновационная или имитационная) более предпочтительна для фирмы в конкретных условиях. Теоретико-игровой анализ такой задачи в контексте технологической конкуренции был предложен в работе [17]. В простейшем варианте рассматривались две фирмы, каждая из которых имеет две стратегии: «разрабатывать новую технологию сейчас» ( $N$ ) или «разрабатывать позже» ( $L$ ), а платежи фирм распределяются следующим образом:

	$L$	$N$
$L$	(8, 8)	(5, 10)
$N$	(10, 5)	(4, 4)

Это симметричная игра, в которой ни у одного из игроков нет преимуществ. В игре два равновесия Нэша в чистых стратегиях ( $(N, L)$  и  $(L, N)$ , причем для фирм они имеют разную ценность) и одно равновесие в смешанных стратегиях ( $x^* = (1/3, 2/3)$ ,  $y^* = (1/3, 2/3)$ ). Возможные варианты поведения, вытекающие из этого анализа: выход на кооперативное поведение с использованием механизмов компенсации или посредника, ожидание в ситуации  $(L, L)$  и отслеживание действий конкурента, переходя в случае необходимости к «стратегии наказания» ( $N$ ), отсрочка старта инновационной активности на  $1/3$  от первоначально рассматриваемого периода и т. д.

Три варианта инновационной стратегии (независимая инновационная деятельность, имитация или создание инновационного консорциума) рассмотрены в статье [16] на основе игры Штакельберга. Простые статические и повторяющиеся игры двух фирм-конкурентов представлены в статье [18] для иллюстрации проблемы о вступлении или невступлении в «инновационную войну». С. Мудамби [19] моделирует инновационную стратегию фирмы, конкурирующей с более сильной компанией, и рассматривает три базовых этапа инновационной деятельности: генерация и разработка идеи, техническая реализация проекта и вывод нового продукта на рынок. Предполагается, что чем раньше сильный конкурент «вмешивается» в этот процесс на любом этапе, тем сильнее ухудшаются результаты инновационного проекта для фирмы. Инновации фирмы могут быть эффективны, если конкурент сочтет слишком затратным или низкодоходным имитацию инновации. Инновационной фирме следует избегать конфронтации с доминирующим на рынке конкурентом, и оптимальная стратегия фирмы, как показывает анализ, заключается в том, чтобы воздерживаться от значительных инвестиций в инновации, которые могут быть имитированы другими компаниями или независимо получены более сильным конкурентом. В пос-

леднем случае для фирмы более выгодна «имитационная» стратегия.

Подобная проблема рассматривается в работе [20], в которой анализируются варианты выбора в условиях дуополии между «радикальной» инновацией (проект  $A$ , направленный на создание принципиально нового продукта) и инкрементальной инновацией (проект  $B$ , направленный на модификацию существующего продукта). Исследование показывает, как характеристики фирм-конкурентов, степень взаимосвязанности продуктов  $A$  и  $B$  и «право первого хода» влияет на принимаемые решения с учетом конъюнктурной неопределенности (спрос) и технической неопределенности (время выполнения проекта). С помощью численных экспериментов показано, что выбор стратегий определяется в основном эффективностью осуществления фирмами проектов разного типа (в том числе их накопленным опытом и знаниями), в то время как «право первого хода» в данной ситуации не оказывает решающего влияния на выбор проектов, хотя и усиливает относительные преимущества фирм.

Более сложный случай выбора стратегий инновационного развития рассмотрен в работе [21], где сравниваются следующие стратегии фирмы, разработавшей новую технологию: использование новой технологии для совершенствования выпускаемого продукта; полная или частичная передача прав на ее использование конкуренту; использование технологии для производства нового продукта и его предложение на рынке вместе с уже выпускаемым. На основе теоретико-игровой модели обсуждаются условия и факторы, влияющие на выбор одной из стратегий.

Ряд теоретико-игровых исследований посвящен вопросам коммерциализации новых разработок. В частности необходимо определить оптимальную стратегию вывода нового продукта, учитывая, прежде всего, существующий потенциал «старого» продукта, далее, необходимость вывода инновации раньше конкурентов, затем, возможные попытки конкурентов имитировать новый продукт и т. д. Подобные задачи решаются на основе конструирования двухэтапных игровых моделей [6, 22–24].

Например, Д. Пурохит [6] предложил модель, в которой на первом этапе фирма-инноватор продает «старый» продукт, а на втором рассматривает возможность вывода «нового» продукта. Конкурент может имитировать «старый» продукт на первом этапе и продавать его на втором. В этих условиях цель инноватора — максимизация прибыли посредством выбора объема продаж «старого» продукта и уровня инновационной активности по раз-



работке «нового» продукта на первом этапе и объема продаж «нового» и «старого» продуктов в условиях конкуренции фирмами-имитаторами на втором этапе. В свою очередь, фирмы-имитаторы на первом этапе принимают решение об имитации «старого» продукта, а на втором — об определении оптимальных объемов его продаж. Рассматриваются несколько стратегий вывода нового продукта (полная замена старого продукта новым, расширение ассортимента (продажа одновременно двух продуктов), постепенная замена старого продукта на новый) и условия их применения [6].

### 3. ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНЫХ ВАРИАНТОВ ПЕРЕДАЧИ ПРАВ НА ОБЪЕКТЫ ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

Для стимулирования фирм, предпринимателей и исследовательских центров вкладывать средства в инновационные проекты применяются различные механизмы защиты прав на объекты интеллектуальной собственности, получаемые в результате инновационной деятельности. Передача прав на использование таких объектов другой фирме (продажа патента или «ноу-хау») может оказаться для их владельца более привлекательным с коммерческой точки зрения, чем самостоятельное коммерческое освоение. Н. Галлини и Р. Винтер [25] на основе теоретико-игровой модели показали, что возможность лицензирования стимулирует инновационную активность, если первоначально применяемые фирмами технологии близки в смысле производственных затрат, но при достаточно сильно различающихся затратах лицензирование, напротив, сдерживает инновационную активность обеих фирм и может приводить к более высоким рыночным ценам на производимую продукцию. Проблема выбора оптимальных вариантов передачи прав на объекты интеллектуальной собственности (выдачи лицензии на их использование (лицензирование), англ. — *licensing*) является одной из центральных тем в работах, посвященных теоретико-игровому исследованию инноваций.

Большая часть исследователей решает эту проблему на основе динамической некооперативной игровой модели [26—29], хотя кооперативный подход также применяется [30, 31]. Но консенсус относительно оптимальных схем лицензирования отсутствует, поскольку выбираемые стратегии весьма чувствительны к ситуативному контексту и исходным предположениям, на основе которых разрабатываются предлагаемые модели. Например, было показано, что в условиях олигополии Курно лицензирование по схеме роялти обеспечивает худший результат для лицензиара, чем лицензи-

рование по схеме фиксированных (паушальных) платежей или при продаже лицензии на аукционе, независимо от объема производства и уровня инновации [26]. М. Камиен и Я. Тауман [32] предложили теоретико-игровую модель, в которой лицензиар выступает в роли лидера в игре Штакельберга, а  $n$  фирм-конкурентов принимают решение о приобретении лицензии. На основе этой модели также было показано, что паушальные платежи предпочтительнее роялти как для лицензиара, так и для потребителей на рынке.

Однако в реальной практике лицензирования паушальные платежи «в чистом виде» применяются сравнительно редко, в основном в случаях, когда передача прав на объект интеллектуальной собственности связана с дополнительными затратами лицензиара (например, по обучению персонала лицензиата) либо когда применение схемы роялти практически невозможно (например, невозможно отслеживать объем продаж). Большинство лицензионных контрактов предусматривают условие выплаты роялти либо «в чистом виде», либо в комбинации с фиксированными платежами [26].

Такой диссонанс между теоретическими выводами и практическими решениями объясняется асимметрией информации, имеющейся у лицензиара и лицензиата, различиями в видах и уровнях инноваций, степенью эксклюзивности передаваемых прав, дифференциацией товаров, переоценкой риска и стремлением его минимизировать, видом конкуренции и другими факторами [28, 33]. Показано, в частности, что если лицензиар и лицензиат конкурируют и их продукты достаточно сильно взаимозаменяемые, то оптимальная схема лицензирования роялти с продаж [34]. При увеличении дифференциации и соответственно, уменьшении конкуренции лицензиара и лицензиата в лицензионном контракте схема роялти будет комбинироваться с фиксированными платежами. В предельном случае, когда продукты полностью различаются, лицензиару следует предлагать лицензиату схему паушального платежа.

Сложности применения теоретических выводов на практике связаны в значительной степени с тем обстоятельством, что теоретико-игровые модели предлагают оптимальный вариант лицензирования с точки зрения лицензиара как игрока, «делающего первый ход» в динамической игре, но не потенциального лицензиата. Однако при заключении лицензионного контракта лицензиат стремится зафиксировать схему платежей, более выгодную для него, и такой схемой, во многих случаях, как раз и оказывается схема роялти. М. Кац и К. Шапиро [26] предложили несколько причин стремления лицензиата к расчету по схеме роялти. Эта схе-

ма является формой разделения риска и одним из вариантов решения проблемы асимметричной информации о ценности инновации, на использование которой передаются права. Если лицензиат по разным причинам не сможет успешно применить приобретаемый патент, передаваемую технологию, «ноу-хау» и др., его выплаты лицензиару будут сравнительно небольшими в случае роялти и, соответственно, затраты лицензиата на приобретение лицензии будут меньше, чем в случае паушального платежа. Подобная мотивация потенциального лицензиата во многих случаях не учитывается при разработке моделей лицензирования.

В качестве примера рассмотрим модель, предложенную в статье [35], включающую в себя несколько вариантов схем оплаты с учетом неопределенности спроса и затрат. Предполагается, что цена на новый продукт  $p(q, w)$ , зависит от объема предложения  $q$  и случайных факторов («состояния природы»)  $w$ . Удельные затраты также рассматриваются как случайная переменная  $c(w)$ . Лицензионный контракт определяет: фиксированный платеж ( $A$ ); роялти с продаж продукции ( $r$ ); роялти с дохода ( $t$ ),  $A > 0$ ,  $r > 0$ ,  $t > 0$ . Конструируемая игра разыгрывается в следующей последовательности:

1) лицензиар предлагает контракт с параметрами  $(t, r, A)$ ;

2) потенциальный лицензиат принимает либо отклоняет предложение лицензиара (в последнем случае игра завершается);

3) лицензиат наблюдает «состояние природы» и принимает решение об объеме выпускаемой продукции.

Игру задают платежные функции игроков, т. е. прибыль лицензиата

$$\pi = (1 - t)p(q, w)q - [c(w) + r]q - A$$

и доход лицензиара от продажи лицензии

$$T(t, r, A; w) = tp[q(t, r; w), w]q(t, r; w) + rq(t, r; w) + A.$$

С помощью этой модели показано, что источник неопределенности оказывает существенное влияние на выбор механизмов оплаты. В частности, при значительной неопределенности спроса выбирается комбинация фиксированных платежей и роялти с дохода. При неопределенности затрат лицензионный контракт должен включать в себя комбинацию фиксированных платежей с одним из вариантов роялти. Включение двух видов роялти в контракт не является оптимальным решением для лицензиата [35].

Д. Сен [28] построил модель, показывающую, как число конкурирующих фирм в условиях олигополии Курно влияет на выбор схемы лицензи-

рования. Согласно этой модели, роялти будет доминировать паушальные платежи и аукционную продажу лицензии при достижении некоторого порогового уровня конкуренции. Оптимальная комбинация фиксированных платежей и роялти рассмотрена в статье [36] с точки зрения «внешнего» инноватора, не участвующего в олигополистической конкуренции на рынке, и «внутреннего» инноватора, который сам конкурирует с другими фирмами на этом рынке. Показано, что «внешний» инноватор будет продавать лицензию всем фирмам-конкурентам, обеспечивая полную диффузию инноваций, и это будет выгодно покупателям, так как будет способствовать росту общего благосостояния, но менее выгодно фирмам-конкурентам. «Внешний» инноватор будет больше инвестировать в развитие технологических инноваций, чем «внутренний» и назначать меньший размер роялти. У обоих инноваторов больше стимулов к инвестированию инновационных разработок в условиях совершенной конкуренции, чем в условиях монополии или дуополии.

М. Кац и К. Шапиро [26] сравнили оптимальные стратегии лицензирования для независимого инноватора (например, университетского исследовательского центра) и инноватора, связанного определенными соглашениями с одной или несколькими фирмами, действующими на рынке и конкурирующими с другими фирмами — потенциальными лицензиатами (случай консорциума). По сравнению с независимым инноватором, консорциум может стремиться ограничить разработку и распространение инноваций. Чем больше фирм участвуют в консорциуме, тем жестче могут быть эти ограничения.

#### 4. МОДЕЛИРОВАНИЕ ВЗАИМОДЕЙСТВИЯ УЧАСТНИКОВ ИННОВАЦИОННОГО КОНСОРЦИУМА

Заметная часть работ по теоретико-игровому анализу инноваций связана с рассмотрением вопросов создания и функционирования консорциумов для совместного осуществления инновационных проектов. Основное место в этих публикациях уделяется анализу проблем стабильности партнерских отношений при совместной инновационной деятельности [37—42].

Многие авторы приходят к выводу, что взаимодействие партнеров строится по принципу прототипной игры «Дилемма узников». Например, в работе [43] проанализированы барьеры для инновационной деятельности в режиме аутсорсинга. Фирма  $A$ , вступая во взаимодействие с фирмой  $B$ , надеется на уменьшение расходов, но в то же время старается передать как можно меньше своих



разработок, знаний и технологий фирме *B*, которая, в свою очередь, стремится к увеличению своей прибыли благодаря взаимодействию с фирмой *A* и старается получить и освоить как можно больше переданных разработок, чтобы увеличить собственную конкурентоспособность в будущем. На примере простой игры авторы показывают, что такой альянс оказывается неустойчивым, и для поддержания его стабильности необходимо привлечь дополнительные механизмы (передача прав на объекты интеллектуальной собственности, создание совместного предприятия и др.). Модель динамического взаимодействия участников подобных альянсов представлена в статье [44] на основе повторяющейся игры «Дилемма узников». Показано, что механизм «жесткой» стратегии наказания «зуб за зуб» недостаточен для эффективного функционирования альянса, и его участники должны выходить на обязывающие и контролируемые партнерские соглашения.

Важный аспект обеспечения стабильности и эффективной работы инновационного консорциума заключается в выборе оптимальных механизмов распределения результатов совместно осуществляемых проектов. Этот вопрос обычно решается на основе распределения суммарного дохода, полученного от проекта, между участниками с учетом их реального вклада. Предложено и используется на практике довольно много вариантов такого распределения: от простых схем распределения пропорционально затратам каждого участника до более сложных моделей, разработанных на основе теории кооперативных игр с трансферабельной полезностью [45–47].

## 5. АНАЛИЗ ИННОВАЦИЙ В СОЦИОЭКОНОМИЧЕСКОМ КОНТЕКСТЕ

Еще одно четко выделяемое направление теоретико-игрового исследования инноваций связано с аспектами деятельности инновационной фирмы в социоэкономическом контексте и государственного регулирования этой деятельности (разработка эффективной инновационной политики, технологическая конкуренция, диффузия инноваций и социальное благосостояние, взаимодействие инновационных фирм и социальных институтов и др.).

Отношения между фирмой-инноватором и покупателями, принимающими и потребляющими инновацию, часто анализируются на основе игры Штакельберга [48]. Игры «инноватор — покупатель» с совершенной информацией и несовершенной информацией рассмотрены в статье [49]. Для анализа поведения потребителей, выбирающих один из нескольких конкурирующих продуктов

или технологий (например, IBM PC и «Macintosh», VHS и «Betamax») и распространения этой технологии разработана игровая модель [50]. В модели, предложенной в работе [51], государство, рассматриваемое как лидер в игре Штакельберга, стремится стимулировать распространение инноваций на рынке, предлагая дотации непосредственно покупателям, приобретающим инновационные продукты, а фирма-инноватор (второй игрок) определяет в этих условиях свои производственные, рекламные и ценовые стратегии. В статье [52] предложено несколько простых игровых моделей для анализа взаимодействия между государственным или негосударственным фондом, финансирующим и «приобретающим» инновационные разработки, и инноватором (индивидуальным исследователем или фирмой), реагирующим на предложение фонда изменением уровня инновационной активности.

Взаимодействие инноватора и государственно-го института, ответственного за разработку инновационной политики, рассматривается в основном в контексте вопросов о защите прав на интеллектуальную собственность. В статье [53] рассмотрено функционирование рынка новых идей и технологий при государственном контроле и без него. На основе теоретико-игровой модели авторы пришли к выводу, что полный отказ от патентов и других форм защиты интеллектуальной собственности способствовал бы повышению общего социального благосостояния. С этих позиций более эффективными стимулами инновационной активности по сравнению с государственной защитой монополии на инновационный продукт оказываются такие экономически проверенные механизмы, как государственные субсидии, премии и гранты, а также антимонопольное регулирование рынка новых идей, подобно тому, как оно осуществляется на рынках товаров и услуг. С другой стороны, как обосновывается с помощью теоретико-игровой модели, представленной в работе [54], фирмы-инноваторы всегда предпочитают режим полной защиты прав интеллектуальной собственности; и ослабление этого режима (например, предоставление возможности свободного некоммерческого использования запатентованной технологии для проведения исследований) не будет стимулировать дорогостоящие инновационные разработки.

С помощью динамической игры *n*-лиц Ф. Шмидт [55] анализирует взаимодействие производителей инновационных продуктов в экономическом контексте, определяемом государственным институтом, организующим «конкурс инноваций», победитель в котором (самый активный инноватор) получит от государства временную эксклюзивную лицензию на работу на данном рынке. Фирмы

принимают решение об инвестировании инноваций и участии в конкурсе. Стратегическими переменными фирм выступают уровень инновационной активности (объем инвестиций в инновации) и объем производства инновационного продукта, определяющий «монопольную ренту», а стратегической переменной государства — временной период, в течение которого предоставляется монопольное право на работу на рынке.

### ЗАКЛЮЧЕНИЕ

Проведенное исследование показывает, что в литературе представлено большое число теоретико-игровых моделей для анализа многих важных аспектов инновационной деятельности и, судя по тенденции роста числа публикаций по этой тематике, в следующем десятилетии можно ожидать существенного повышения исследовательской активности в этой области. Однако доля исследований по теме инноваций, проводимых с помощью методов теории игр, на сегодня ничтожно мала по сравнению с общим объемом публикаций по теме инновационной деятельности. В общем объеме работ по теории игр число исследований, связанных с инновациями, также пока весьма мало. Кроме того, многие из предлагаемых моделей оказываются либо слишком простыми, что не позволяет их использовать для принятия реальных решений, либо слишком абстрактными и формальными, что также затрудняет их практическое применение. Однако в публикациях двух последних десятилетий наблюдается четкая тенденция уточнения и конкретизации базовых формальных теоретико-игровых моделей и стремление обеспечить практическую применимость разрабатываемых моделей. Также обращает на себя внимание заметно увеличившееся за последние 10–15 лет число попыток верификации и обоснования предложенных моделей с помощью машинных экспериментов и лабораторных опытов. Заслуживает внимания также тот факт, что работы по рассматриваемой теме в последние годы публикуются не только в специальных академических журналах по теории игр («International Game Theory Review», «Games and Economic Behavior», «International Journal of Game Theory» и др.) и математически ориентированных экономических журналах, но и во многих прикладных экономико-управленческих и бизнес-изданиях. С учетом выделенных тенденций развития теоретико-игровых исследований инновационной деятельности, широкого спектра моделируемых проблем и мощного логико-математического инструментария, теоретико-игровое направление представляется весьма перспективным для исследова-

вания и анализа проблем, связанных с инновационным развитием как отдельных фирм, так и целых регионов и государств.

### ЛИТЕРАТУРА

1. *Beath J., Katsoulacos Y., Ulph D.* The game-theoretic analysis of innovation: a survey // *Bulletin of Economic Research*. — 1989. — Vol. 41, N 3. — P. 163–184.
2. *Reinganum J.F.* Practical implications of game theoretic models of R&D // *The American Economic Review*. — 1984. — Vol. 74, N 2. — P. 61–66.
3. *Yang H.-L., Wu T.C.T.* Knowledge sharing in an organization // *Technological Forecasting and Social Change*. — 2008. — Vol. 75. — P. 1128–1156.
4. *Новиков Д.А., Иващенко А.А.* Модели и методы организационного управления инновационным развитием фирмы. — М.: КомКнига, 2006. — 332 с.
5. *Plambeck E.L., Taylor T.A.* Sell the plant? The impact of contract manufacturing on innovation, capacity, and profitability // *Management Science*. — 2005. — Vol. 51, N 1. — P. 133–150.
6. *Purohit D.* What should you do when your competitors send in the clones? // *Marketing Science*. — 1994. — Vol. 13, N 4. — P. 392–411.
7. *Park W.G.* A note on innovation and patent protection: intertemporal imitation-risk smoothing // *Economics Letters*. — 1997. — Vol. 54. — P. 185–189.
8. *Dasgupta P., Stiglitz J.E.* Industrial structure and the nature of innovative activity // *The Economic Journal*. — 1980. — Vol. 90, N 358. — P. 266–293.
9. *Dasgupta P., Stiglitz J.E.* Uncertainty, industrial structure and the speed of R&D // *Bell Journal of Economics*. — 1980. — Vol. 11, N 1. — P. 1–28.
10. *Cellini R., Lambertini L.* Private and social incentives towards investment in product differentiation // *International Game Theory Review*. — 2004. — Vol. 6, N 4. — P. 493–508.
11. *Smit H.T.J., Trigeorgis L.* Real options and games: competition, alliances and other applications of valuation and strategy // *Review of Financial Economics*. — 2006. — Vol. 15. — P. 95–112.
12. *Smit H.T.J., Trigeorgis L.* Strategic options and games in analysing dynamic technology investments // *Long Range Planning*. — 2007. — Vol. 40. — P. 84–114.
13. *Amaldoss W., Sanjay J.* David vs. Goliath: an analysis of asymmetric mixed-strategy games and experimental evidence // *Management Science*. — 2002. — Vol. 48, N 8. — P. 972–991.
14. *Petit M.-L., Sanna-Randaccio F., Tolwinski B.* Innovation and foreign investment in a dynamic oligopoly // *International Game Theory Review*. — 2000. — Vol. 2, N 1. — P. 1–28.
15. *Bischi G.-I., Lamantia F.A.* Competition game with knowledge accumulation and spillovers // *International Game Theory Review*. — 2004. — Vol. 6, N 3. — P. 323–341.
16. *Chen S., Cheng B., Shao Y.* Studies on the choice of enterprise's innovation strategies under two market circumstances // *International Management Review*. — 2007. — Vol. 3, N 4. — P. 63–68.
17. *Dasgupta P.* The theory of technological competition // *Economic Organizations as Games*. — Oxford: Basil Blackwell, 1989. — P. 139–164.
18. *Zhang R.* Game analysis on enterprise technological innovation // *China-USA Business Review*. — 2007. — Vol. 6, N 3. — P. 63–66.
19. *Mudambi S.M.* The games retailers play // *Journal of Marketing Management*. — 1996. — Vol. 12. — P. 695–706.



20. *Ali A., Kalwani M., Kovenock D.* Product development projects: pioneering versus incremental innovation strategies // *Management Science*. — 1993. Vol. 39, N 3. — P. 255–274.
21. *Costa L.A., Dierckx I.* The strategic deployment of quality-improving innovations // *Journal of Business*. — 2005. — Vol. 78, N 3. — P. 1049–1072.
22. *Sankaranarayanan R.* Innovation and the durable goods monopolist: the optimality of frequent new-version releases // *Marketing Science*. — 2007. — Vol. 26, N 6. — P. 774–791.
23. *Erat S., Kavadias S.* Introduction of new technologies to competing industrial customers // *Management Science*. — 2006. — Vol. 52, N 11. — P. 1675–1688.
24. *Wakeman S.* A dynamic theory of technology commercialization strategy // *Academy of Management Proceedings*. — 2008. — P. 1–6.
25. *Gallini N.T., Winter R.A.* Licensing in the theory of innovation // *The RAND Journal of Economics*. — 1985. — Vol. 16, N 2. — P. 237–252.
26. *Katz M.L., Shapiro C.* How to license intangible property // *Quarterly Journal of Economics*. — 1986. — Vol. 101. — P. 567–589.
27. *Haller H., Pavlopoulos A.* Technological failure, economic success // *International Game Theory Review*. — 2002. — Vol. 4, N 4. — P. 415–434.
28. *Sen D.* Fee versus royalty reconsidered // *Games and Economic Behavior*. — 2005. — Vol. 53. — P. 141–147.
29. *Erutco C., Richelle Y.* Optimal licensing contracts and the value of a patent // *Journal of Economics and Management Strategy*. — 2007. — Vol. 16. — P. 407–436.
30. *Tauman Y., Watanabe N.* The Shapley value of a patent licensing game: the asymptotic equivalence to non-cooperative results // *Economic Theory*. — 2007. — Vol. 30. — P. 135–149.
31. *Jelšov A., Tauman Y.* The private value of a patent: a cooperative approach // *Mathematical Social Sciences*. — 2009. — Vol. 58. — P. 84–97.
32. *Kamien M.I., Tauman Y.* Fees versus royalties and the private value of a patent // *Quarterly Journal of Economics*. — 1986. — Vol. 101, N 3. — P. 471–491.
33. *Gallini N.T., Wright B.D.* Technology transfer under asymmetric information // *The RAND Journal of Economics*. — 1990. — Vol. 21, N 1. — P. 147–160.
34. *Mukherjee A., Balasubramanian N.* Technology transfer in a horizontally differentiated product market // *Research in Economics*. — 2001. — Vol. 55. — P. 257–274.
35. *Bousquet A., Cremer H., Ivaldi M., Wolkowicz M.* Risk sharing in licensing // *International Journal of Industrial Organization*. — 1998. — Vol. 16. — P. 535–554.
36. *Sen D., Tauman Y.* General licensing schemes for a cost-reducing innovation // *Games and Economic Behavior*. — 2007. — Vol. 59. — P. 163–186.
37. *Chen F., Fan L.* Analysis on stability of strategic alliance: a game theory perspective // *Science: A Journal of Zhejiang University*. — 2006. — Vol. 7, N 12. — P. 1995–2001.
38. *Binenbaum E.* Incentive issues in R&D consortia: insights from applied game theory // *Contemporary Economic Policy*. — 2008. — Vol. 26, N 4. — P. 636–650.
39. *Wolters H., Schuller F.* Explaining supplier-buyer partnerships: a dynamic game theory approach, *European Journal of Purchasing and Supply Management*. — 1997. — Vol. 3, N 3. — P. 155–164.
40. *Cousins P.* A conceptual model for managing long-term inter-organisational relationships // *European Journal of Purchasing and Supply Management*. — 2002. — Vol. 8. — P. 71–82.
41. *Brocas I.* Vertical integration and incentives to innovate // *International Journal of Industrial Organization*. — 2003. — Vol. 21. — P. 457–488.
42. *Duvallet J., Garapin A., Llerena D., Robin S.* A game theoretical approach of price negotiation and coordination in an innovative firm-supplier context: an experimental analysis // *International Negotiation*. — 2004. — Vol. 9. — P. 245–269.
43. *Bergfeld M.M., Doepfer B.* Innovation in outsourcing alliances: managing the Prisoner's Dilemma of cooperative competence building // *Proceedings of the XX International Society for Professional Innovation Management Conference — The Future of Innovation*. — Vienna: ISPIM, 2009.
44. *Arend R.J., Seale D.A.* Modeling alliance activity: an iterated Prisoners' Dilemma with exit option // *Strategic Management Journal*. — 2005. — Vol. 26. — P. 1057–1074.
45. *Jarimo T., Pulkkinen U., Salo A.* Encouraging suppliers to process innovations: a game theory approach // *International Journal of Technology Intelligence and Planning*. — 2005. — Vol. 1, N 4. — P. 403–423.
46. *Watanabe N., Muto S.* Stable profit sharing in bargaining outcomes // *International Journal of Game Theory*. — 2008. — Vol. 37. — P. 505–523.
47. *Zakharov V., Gan'kova A., Dementieva M., Neittaanmaki P.* Comparing solutions in joint implementation projects // *International Game Theory Review*. — 2008. — Vol. 10, N 1. — P. 119–128.
48. *Wu D., Erkoç M., Karabuk S.* Managing capacity in the high-tech industry: a review of literature // *The Engineering Economist*. — 2005. — Vol. 50. — P. 125–158.
49. *Grimaldi G.* Creating a “market for inventions”: a referee model // *Journal of Technol. Manag. Innovation*. — 2008. — Vol. 3, N 3. — P. 33–44.
50. *Tomochi M., Murata H., Kono M.* A consumer-based model of competitive diffusion: the multiplicative effects of global and local network externalities // *Journal of Evolutionary Economics*. — 2005. — Vol. 15. — P. 273–295.
51. *De Cesare L., Di Liddo A.* A Stackelberg game of innovation diffusion: pricing, advertising and subsidy strategies // *International Game Theory Review*. — 2001. — Vol. 3, N 4. — P. 325–339.
52. *Devine S., Webb C.* Competitive funding: a game theory analysis of how research organisations adapt their behaviour in the New Zealand competitively funded science system // *Science and Public Policy*. — 2004. Vol. 31, N 5. — P. 407–414.
53. *Boldrin M., Levine D.K.* The economics of ideas and intellectual property // *Proceedings of the National Academy of Sciences of the United States of America*. — 2005. — Vol. 102, N 4. — P. 1252–1256.
54. *Moschini G., Yerokhin O.* Patents, research exemption, and the incentive for sequential innovation // *Journal of Economics and Management Strategy*. — 2008. — Vol. 17, N 2. — P. 379–412.
55. *Schmidt F.* Innovation contests with temporary and endogenous monopoly rents // *Rev. Econ. Design*. — 2008. — Vol. 12. — P. 189–208.

Статья представлена к публикации членом редколлегии Р.М. Нижнегородцевым.

**Дубина Игорь Николаевич** — канд. филос. наук, докторант кафедры теоретической кибернетики и прикладной математики, Алтайский государственный университет, ☎ (385-2) 24-65-58, ✉ din@asu.ru.

# РЕШЕНИЕ ОБОБЩЕННОЙ ЗАДАЧИ ДЖОНСОНА С ОГРАНИЧЕНИЯМИ НА СРОКИ ВЫПОЛНЕНИЯ ЗАДАНИЙ И ВРЕМЕНА РАБОТЫ МАШИН. Ч. 2. Приближенные методы решения

Ю.А. Зак

Разработан комплексный алгоритм решения flow-shop problem, использующий основанные на различных гипотезах эвристики и правила предпочтения. Предложены правила построения последовательностей выполнения заданий, позволяющие анализировать различные области притяжения локальных минимумов. Показано, что эффективность построенных расписаний повышается благодаря предложенному в работе алгоритму локальной вариации. Дано сравнение с ранее полученными решениями, приведен числовой пример.

**Ключевые слова:** оптимальное расписание, последовательность выполнения заданий, эвристики, локальные вариации.

## ВВЕДЕНИЕ

В первой части работы [1] была приведена постановка задачи, предложены и обоснованы расчетные формулы для вычисления нижней границы суммарной длины расписания и разработаны точные и приближенные алгоритмы решения задачи на основе метода «ветвей и границ». Вторая часть посвящена развитию приближенных методов решения задачи в традиционной постановке (без учета ограничений на допустимые сроки выполнения заданий и ресурсы работы рабочих станций) в условиях большой размерности.

Еще в 1964 г. Джонсоном [2, 3] было доказано, что для числа рабочих станций  $m = 2$  может быть получено точное решение и простое решающее правило сформулированной задачи, заключающееся в следующем. Все множество заданий  $I = \{1, 2, \dots, i, \dots, n\}$  может быть разбито на два подмножества  $I_1 = \{i : t(i, 1) \leq t(i, 2)\}$  и  $I_2 = \{i : t(i, 1) > t(i, 2)\}$ , где  $t(i, 1)$  и  $t(i, 2)$  — соответственно длительности первой и второй операций (работ)  $i$ -го здания. Если все задания подмножества  $I_1$  расположены в порядке возрастания, а подмножества  $I_2$  в порядке убывания длительностей операций, то в оптималь-

ной последовательности выполнения заданий вначале следует упорядоченное подмножество заданий  $I_1$ , а затем — упорядоченное подмножество  $I_2$ .

Если число рабочих станций больше двух, то сформулированное Джонсоном [2, 3] и описанное в первой части настоящей работы простое решающее правило уже не справедливо, однако оно может быть использовано для построения различных эвристик при конструировании различных алгоритмов получения приближенных решений.

На основе большого вычислительного опыта успешного решения многих прикладных и тестовых задач (см., например, работы [4–6]) могут быть сформулированы некоторые эвристические принципы построения эффективных последовательностей выполнения заданий, в которых в первую очередь должны выполняться задания, у которых:

- «узкое место» (операция с максимальной продолжительностью) расположена дальше всех от первой станции;
- первая операция имеет самую короткую продолжительность (предпочтение первой операции);
- суммарное время всех операций, исключая первую, наиболее длинное (предпочтение оставшейся длинной операции).



Сформулированные эвристические положения лежат в основе многих эвристических, статистических и генетических алгоритмов решения данного класса задач в условиях большой размерности. Эффективные формулы вычисления нижних границ оптимального решения, приведенные в первой части работы, позволяют эффективно останавливать процесс решения при получении последовательности выполнения заданий с суммарной длительностью, удовлетворяющей поставленным требованиям точности.

Многочисленные эвристические методы и правила предпочтения [2, 4–16], генетические алгоритмы (см., например, работы [10, 17–19]) позволяют зачастую найти решение, принадлежащее области глобального или эффективного локального минимума. Однако никакая из предложенных эвристик не гарантирует попадание в область глобального минимума. В результате локального улучшения последовательности выполнения заданий можно в ряде случаев повысить эффективность и сократить длину построенного расписания. Предлагаемые в литературе правила локального улучшения последовательности выполнения заданий путем обмена местами двух индексов [2, 7, 8, 10, 17] носят либо локальный, либо статистический характер и потому не могут быть интерпретированы как методы локального спуска в область глобального минимума. С этой точки зрения представляет интерес создание гибридных методов, которые на основе хороших эвристик позволяют попасть в зону глобального минимума, где методами локальных перестановок местами заданий осуществляется спуск в точку локального минимума этой области. Принимая во внимание полиномиальную сложность и сравнительно небольшой объем вычислений каждого из эвристических алгоритмов, представляет интерес в таком комплексном алгоритме использовать сразу несколько наиболее эффективных по результатам вычислительных экспериментов [9, 10, 17] эвристик.

В работе предложены некоторые гибридные алгоритмы и приближенные методы решения задач большой размерности, использующие различные известные ранее эвристики и правила предпочтения, а также локальные улучшения построенных на основе эвристик последовательностей. Как показали вычислительные эксперименты, в ряде случаев предлагаемые методы позволяют повысить эффективность получаемых известными алгоритмами приближенных решений при незначительном объеме вычислений.

Предложенные и исследованные в работе модификации метода локальной вариации построенного допустимого решения в ряде случаев позволяют

улучшить построенную допустимую последовательность выполнения заданий и могут найти применение в различных эвристических, статистических и генетических алгоритмах решения задачи.

В первой части работы была введена система обозначений, приведены постановка и математическая формулировка задачи, а также формульные выражения для оценки нижней границы длины строящегося расписания. В дальнейшем изложении используются введенные ранее обозначения, установленные свойства и формулы.

Нумерация формул второй части работы продолжает нумерацию формул первой части.

### 1. ЛОКАЛЬНЫЕ ПРИОРИТЕТНЫЕ ПРАВИЛА ДЛЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ВЫПОЛНЕНИЯ ЗАДАНИЙ

Далее исследуются локальные свойства последовательностей выполнения заданий, позволяющие осуществить обмен местами двух рядом стоящих заданий, гарантируя выполнение системы ограничений и обеспечение значения критерия оптимальности не хуже, чем в исходной последовательности.

Рассмотрим некоторый частичный план  $P^s$ , в котором определена подпоследовательность выполнения заданий  $\tilde{J}^s = \{i_1^s, i_2^s, \dots, i_{l-1}^s, i_l^s, \dots, i_R^s\}$  и рассчитаны значения  $\theta^s(i, k)$  и  $\sigma^s(i, k)$  всех операций  $(l, k)$  подпоследовательности  $\tilde{J}^s$ . Рассмотрим некоторую другую подпоследовательность, в которой задания  $i_{l-1}^s$  и  $i_l^s$  поменяны местами. Обозначим соответственно  $\theta^s(i, k | i_{l-1}^s, i_l^s)$ ,  $\sigma^s(i, k | i_{l-1}^s, i_l^s)$  и  $\theta^s(i, k | i_l^s, i_{l-1}^s)$ ,  $\sigma^s(i, k | i_l^s, i_{l-1}^s)$  времена начала и окончания выполнения всех операций для этих двух подпоследовательностей  $\tilde{J}^s(i_{l-1}^s, i_l^s)$  и  $\tilde{J}^s(i_l^s, i_{l-1}^s)$ .

**Утверждение 1.** Если для пары заданий  $i_{l-1}^s$  и  $i_l^s$  некоторой подпоследовательности  $\tilde{J}^s$  выполняется равенство  $\sigma^s(i_l^s, m) = \sigma^s(i_{l-1}^s, m) + t(i_l^s, m)$ , то обмен местами этих заданий в подпоследовательности  $\tilde{J}^s$  не может привести к сокращению суммарного времени выполнения всех заданий (локальному минимуму), и потому в смысле локальных вариаций рассматривается как нецелесообразный. ♦

Утверждение 1 позволяет выделить в последовательности только те пары заданий, для которых

$$\sigma^s(i_l^s, m) > \sigma^s(i_{l-1}^s, m) + t(i_l^s, m), \quad (26)$$

и обмен местами которых в данной перестановке может оказаться целесообразным.

Следующие далее утверждения позволяют проверить локальную эффективность такой «перспективной» перестановки, зачастую не производя расчета времен завершения выполнения всех заданий, стоящих за этой парой в последовательности.

**Утверждение 2 (правило абсолютного предпочтения).** Если справедлива система неравенств

$$\begin{aligned} \sigma^s(i_{l-1}^s, k|l, l-1) &\leq \sigma^s(i_l^s, k|l-1, l), \\ k &= 1, \dots, m, \end{aligned} \quad (27)$$

то подпоследовательность  $\tilde{J}^s(i_l^s, i_{l-1}^s)$  имеет абсолютное предпочтение перед подпоследовательностью  $\tilde{J}^s(i_l^s, i_{l-1}^s)$ . Если подпоследовательность  $\tilde{J}^s$  содержит допустимые расписания выполнения заданий, то перестановка в ней местами двух рядом стоящих заданий  $i_{l-1}^s$  и  $i_l^s$  не приводит к исключению допустимых или более перспективных и оптимальных расписаний. ♦

Если для рассматриваемых подпоследовательностей справедливо неравенство

$$\sigma^s(i_{l-1}^s, m|l, l-1) > \sigma^s(i_l^s, m|l-1, l), \quad (28)$$

то перестановка местами этих двух индексов может привести к увеличению длины расписания или нарушению системы ограничений.

**Правила относительного предпочтения.** Обозначим  $K$  — множество всех рабочих станций, а  $K_1$  — некоторое его подмножество. Если для подпоследовательностей  $\tilde{J}^s(i_{l-1}^s, i_l^s)$  и  $\tilde{J}^s(i_l^s, i_{l-1}^s)$  справедлива система неравенств

$$\begin{aligned} \sigma^s(i_{l-1}^s, m|l, l-1) &\leq \sigma^s(i_l^s, m|l-1, l), \\ \sigma^s(i_{l-1}^s, k|l, l-1) &\leq \sigma^s(i_l^s, k|l-1, l), \\ k &\in K_1; \end{aligned} \quad (29)$$

$$\sigma^s(i_{l-1}^s, k|l, l-1) > \sigma^s(i_l^s, k|l-1, l), \quad k \in K/K_1, k \neq m; \quad (30)$$

$$\begin{aligned} \theta^s(i_{l-1}^s, k|l, l-1) + \sum_{i \in I_2^s} t(i, k) &\leq H_k, \\ k &= 1, \dots, m, \end{aligned} \quad (31)$$

то подпоследовательность  $\tilde{J}^s(i_l^s, i_{l-1}^s)$  имеет относительное предпочтение перед подпоследовательностью  $\tilde{J}^s(i_l^s, i_{l-1}^s)$ .

В случае относительного предпочтения перестановка двух рядом стоящих индексов может

привести к нарушению ограничений на завершение выполнения отдельных заданий и на времена работы машин, а также к увеличению длины расписания выполнения всего комплекса работ. Поэтому для окончательных выводов о целесообразности перестановки местами этих двух индексов необходимо рассчитать сроки завершения и следующих заданий, стоявших в исходной подпоследовательности  $\tilde{J}^s$  после задания  $i_l^s$ .

**Утверждение 3.** Если для некоторого задания  $j$ , стоявшего в подпоследовательности  $\tilde{J}^s$  после задания  $i_l^s$ , справедливы правила абсолютного предпочтения в виде

$$\begin{aligned} \sigma^s(j, k|l, l-1) &\leq \sigma^s(j, k|l-1, l), \\ k &= 1, \dots, m, \end{aligned} \quad (32)$$

то обмен местами индексов  $i_{l-1}^s$  и  $i_l^s$  не приведет к нарушению системы ограничений на времена работы рабочих систем или к увеличению длины расписания.

Если для какого-либо задания  $j$ , стоящего в подпоследовательности  $\tilde{J}^s$  после задания  $i_l^s$ , условия (31) не выполняются для последней операции задания (для  $k = m$ ), то обмен местами индексов  $i_{l-1}^s$  и  $i_l^s$  нецелесообразен. ♦

Так как пересчет сроков завершения выполнения всех операций заданий, стоящих в подпоследовательности  $\tilde{J}^s$  после задания  $i_l^s$ , связан с определенным объемом вычислений, для практических приложений рекомендуется анализировать не более 2—3 следующих за заданием  $i_l^s$  заданий.

## 2. АЛГОРИТМ ЛОКАЛЬНОЙ ВАРИАЦИИ ПОСЛЕДОВАТЕЛЬНОСТИ ВЫПОЛНЕНИЯ ЗАДАНИЙ (LV-АЛГОРИТМ)

Пусть построена некоторая допустимая последовательность выполнения всего множества заданий  $\tilde{L} = \{i_1, i_2, \dots, i_n\}$ . Рассмотрим, как на основе локальных перестановок местами двух рядом стоящих в последовательности индексов сократить длину расписания выполнения всего комплекса работ. Определим следующие параметры алгоритма:

— параметр  $\rho$  определяет число стоящих в последовательности после рассматриваемой пары заданий, которые анализируются в соответствии с правилами абсолютного и относительного предпочтений;

— параметр  $\beta$  определяет стратегию локального поиска: при  $\beta = 0$  обмен местом расположения в последовательности пары индексов производится



до тех пор, пока он потенциально может привести к сокращению длины расписания; при  $\beta = 1$  перестановка местами двух рядом стоящих в последовательности заданий происходит до тех пор, пока процесс обмена не достигнет двух стоящих последними в последовательности  $\tilde{L}$  индексов (одно-разовый просмотр последовательности  $\tilde{L}$ );

— счетчики  $N1$ ,  $N2$  и  $N3$ , определяющие целесообразность продолжения отдельных шагов алгоритма.

LV-алгоритм состоит из следующих шагов.

*Шаг 0 (предварительный).* Устанавливаются  $N1 = N2 = 0$ ,  $N3 = n - 1$ , переход к шагу 1.

*Шаг 1.* Если  $N1 = N3$  и  $N2 = 0$ , то переход к шагу 4. Если  $N1 = N3$  и  $N2 > 0$ , то при  $\beta = 0$  полагается  $M1 = 0$ ,  $N2 = 0$  и переход к шагу 2, а при  $\beta = 1$  — алгоритм заканчивает работу. Если  $N1 < N3$ , то переход к шагу 2.

*Шаг 2.* Если для пары индексов  $i_{l-1}^s = M1$  и  $i_l^s = M1 + 1$  выполняются условия (26), то проверяются для них условия абсолютного предпочтения (32). В случае их выполнения производится обмен местами этих заданий в последовательности, в счетчики  $N1$  и  $N2$  добавляется 1 и переход к шагу 4. Если условия (26) не выполняются, либо при выполнении условия (26) справедливо неравенство (28),  $M1$  увеличивается на единицу и переход к шагу 1. При выполнении условия (26) совместно с правилами относительного предпочтения (29)—(31)  $M1$  увеличивается на единицу и переход к шагу 3.

*Шаг 3.* Определяем число анализируемых заданий, стоявших в последовательности после задания  $i_l^s$ , сроки завершения которых должны быть рассчитаны для определения локальной перестановки этих двух индексов:  $\rho1 = \min(\rho, R - M1 - 1)$ . Для заданий  $j = i_p^s$ ,  $p = l + 1, \dots, l + \rho1$ , подпоследовательности  $J^s(i_l^s, i_l^s - 1)$  вычисляются сроки завершения всех операций  $\sigma^s(j, k|l, l - 1)$ ,  $k = 1, \dots, m$ , и проверяется выполнение условий (27) и (28). Если для какого-либо  $j$  выполняются неравенства (27), то значение счетчика  $N2$  увеличивается на 1 и переход к шагу 1. Если условия (28) выполняются, либо ни для одного из этих  $\rho1$  заданий ни одна из систем неравенств (27) или (28) не выполняется, то переход к шагу 1.

*Шаг 4.* Для всех заданий  $j = i_p^s$ ,  $p = l + 1, \dots, l + R$ , новой подпоследовательности  $J^s(i_l^s, i_{l-1}^s)$  пересчитываются времена завершения операций на каждой машине. Переход к шагу 1.

### 3. ГИБРИДНЫЕ ПРИБЛИЖЕННЫЕ И ЭВРИСТИЧЕСКИЕ МЕТОДЫ РЕШЕНИЯ FLOW-SHOP PROBLEM

Предложенные Джонсоном [3] простые правила предпочтения для получения точного решения в случае двух машин нашли широкое применение при построении различных эвристических приближенных алгоритмов решения задачи в случае, когда число машин  $m > 2$ . Далее на основе предложенных ранее отдельных эвристик конструируется некоторый комплексный алгоритм генерирования, локального улучшения и оценки степени приближения построенных на основе различных эвристик расписаний с последующим отбором наилучшего из них. Построенное расписание может быть в дальнейшем улучшено описанным в § 2 алгоритмом локальных вариаций.

В конструируемом алгоритме используются следующие эвристики и правила предпочтения.

- Эвристика [6], согласно которой все машины разделяются на две группы. Для каждого задания суммируются времена обработки на машинах каждой группы

$$T_1(i, r) = \sum_{k=1}^r t(i, k), \quad T_2(i, r) = \sum_{k=r+1}^m t(i, k), \\ i = 1, \dots, n, \quad r = 0, \dots, m, \quad (33)$$

и задача в дальнейшем решается построением расписания с помощью алгоритма Джонсона [3] для случая двух машин. По результатам решения таких  $(m - 1)$  проблем выбирается наилучшее решение. Сложность такого алгоритма равна  $O(nm^2 + mn \log n)$  [6].

- Предложенные в работе [7] выражения в виде

$$T_1(i) = \sum_{k=1}^m (m - k + 1)t(i, k), \\ T_2(i) = \sum_{k=1}^m kt(i, k) \quad (34)$$

с дальнейшим одноразовым применением алгоритма Джонсона.

- Построение приоритетных правил с присвоением каждому заданию некоторого ранга в соответствии с алгоритмом [20] (engl-slope index), вычисляемого по формуле

$$V(i) = \sum_{k=1}^m (2k - m - 1)t(i, k) = \\ = 2 \sum_{k=1}^m kt(i, k) - (m - 1) \sum_{k=1}^m t(i, k), \quad (35)$$

затем задания располагаются в последовательности в соответствии с монотонно убывающим рангом.

Заметим, что наряду с применяемыми можно воспользоваться и другими правилами ранжирования заданий, предложенных в работах [4, 5] и др. Алгоритм [11] со сложностью  $O(mn^3)$  вытекает из выражения (33) при  $r = 0$ , когда все задания упорядочиваются в соответствии с убыванием значений  $T_2(i, 0)$ .

Отличительные особенности предлагаемого алгоритма:

- возможность анализа эффективности полученного решения на основе вычисленной по формуле (1) или (12) оценки (см. первую часть [1] работы);

- процедура локального улучшения каждого построенного на основе различных эвристик расписания с помощью LV-алгоритма.

Существенные преимущества описанного в § 2 LV-алгоритма относительно известных в литературе (см., например, [2, 7, 8, 10, 17]) методов локальной вариации построенных последовательностей (перестановок) заключаются в:

- целенаправленности выбора рядом стоящей пары индексов для обмена;

- использовании решающих правил для оценки возможности локального улучшения расписания, что требует существенно меньшего объема вычислений;

- достижении локального минимума для рассматриваемой последовательности.

Длина каждого вновь построенного и локально улучшенного расписания сравнивается с оценкой нижней границы оптимального решения, вычисленной в соответствии с выражением (1), и лучшим решением, полученным по результатам проведенных вычислений. При достижении требуемой точности полученного результата процесс решения завершается.

Рассматриваемая задача многоэкстремальная. Какая-то из используемых и построенных из различных соображений эвристик должна обеспечить попадание в некоторую область локального или глобального минимума со значением критерия оптимальности, близким к длине оптимального расписания. Дальнейший спуск в точку локального минимума этой области методом локальных вариаций в ряде случаев должен повысить эффективность полученного эвристическими алгоритмами решения.

### 3.1. Алгоритм В

Алгоритм приближенного решения задачи состоит из следующих шагов.

*Шаг 0 (предварительный).* Полагается  $E = \infty$ ,  $r = 0$ . Устанавливается требуемая точность решения задачи  $\varepsilon \geq 1$  (в частности, если  $\varepsilon = 1$ , ставится требование получение точного решения). Вводится некоторый параметр алгоритма  $Z$ , определяющий логическую последовательность выполнения шагов. Полагается  $Z = 0$ . Рассчитывается нижняя граница длины оптимального расписания по формуле (1), по значению которой определяется  $F_{\min}$ , переход к шагу 1.

*Шаг 1.* Рассчитываются модифицированные времена выполнения каждого из заданий  $T_1(i)$  и  $T_2(i)$ ,  $i = 1, \dots, n$ , в соответствии с формулой (34). Переход к шагу 2.

*Шаг 2.* Строится последовательность выполнения заданий на основе алгоритма Джонсона. Полагается  $F$  — длина построенного расписания. Переход к шагу 3.

*Шаг 3.* Если  $F < E$ , то полагается  $E = F$  и запоминается построенная последовательность как наилучшее из полученных к данному моменту решений. Если  $E \leq \varepsilon E_{\min}$ , то получено решение задачи, и алгоритм завершает работу. В противном случае, если  $Z = 0$ , переход к шагу 4, положив  $Z = 1$ . Если  $Z \neq 0$ , переход к шагу  $(Z + 4)$ .

*Шаг 4.* Выполняется LV-алгоритм локальной вариации последовательности выполнения заданий. Переход к шагу 3.

*Шаг 5.* Расчет рангов выполнения заданий в соответствии с формулой (35), строится последовательность выполнения заданий в соответствии с монотонно убывающим рангом. Полагается  $F$  — длина построенного расписания. Переход к шагу 3, положив  $Z = 2$ .

*Шаг 6.* Если  $r \leq m$ , рассчитываются модифицированные времена выполнения каждого из заданий  $T_1(i)$  и  $T_2(i)$ ,  $i = 1, \dots, n$ , в соответствии с формулой (33). Переход к шагу 2, положив  $r = r + 1$ . Если  $r > m$ , то алгоритм заканчивает работу. Запомненная в памяти последовательность выполнения заданий с длиной расписания равной  $E$  является решением задачи. ♦

Разработано программное обеспечение, позволяющее решать практические задачи размерностью до нескольких десятков и даже сотен заданий и выводить результаты решения как в форме таблиц, так и в графической форме — в виде диаграмм Ганта.

Многочисленные вычислительные эксперименты показали эффективность предлагаемых алгоритмов. Решения с длиной расписания, не превышающей 1–3 % вычисленной по формулам (1) нижней границы, были получены за полиномиальное время. Кроме того, в процессе решения часто



Таблица 1

## Исходные данные примера

$i$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
1	15	18	25	30
2	12	27	30	5
3	25	30	40	18
4	35	5	14	20
5	40	10	5	5
6	10	20	30	10
7	15	35	40	10
8	18	12	30	15
9	20	25	30	25
10	15	18	10	45

удается построить несколько расписаний, удовлетворяющих требованиям точности, длина которых отличается несущественно. Если последовательности выполнения заданий в таких расписаниях совершенно различны, то наиболее приемлемое из них может выбираться из других производственных соображений или согласно другому критерию оптимальности.

## 4. ИЛЛЮСТРАТИВНЫЙ ПРИМЕР

Времена выполнения каждой из четырех операций для 10 заданий представлены в табл. 1. Нижняя граница длины расписания, вычисленная по формуле (1) определяется машиной 4 и равна 289. Решение примера, полученное с помощью алгоритма В, сведено в табл. 2.

Расчитанные на шаге 1 алгоритма А в соответствии с эвристикой [7] модифицированные времена вы-

полнения заданий представлены в столбцах 2 и 3 табл. 2. Последовательность выполнения заданий, рассчитанная для этих времен алгоритмом Джонсона [3] с длиной расписания 304, представлена в столбцах 4–8. Далее на шаге 3 с помощью LV-алгоритма была осуществлена локальная вариация построенной последовательности. Целесообразным был лишь обмен местами заданий 3 и 7, 7 и 2, а также 4 и 2 (см. условия (2)). Как показали правила абсолютного и локального приоритета (3) и (27), (4) и (28), первые два обмена не приводят к более эффективному результату. Обмен местами заданий 4 и 2 привел к расписанию, длина которого равна нижней границе. Следовательно, это расписание оптимально, и необходимость в выполнении дальнейших шагов отпадает.

Предложенный в работе алгоритм локальной вариации уже после выполнения его первого шага позволил существенно (на 15 временных интервалов) повысить эффективность построенного на основе эвристических правил расписания.

Отметим, что построенная в соответствии с правилами предпочтения (10) [20] последовательность выполнения заданий  $\Pi = \{10, 1, 9, 6, 8, 7, 3, 2, 4, 5\}$  с длиной расписания  $F = 315$ , которая, будучи улучшена алгоритмами локальной вариации, имеет вид  $\Pi = \{10, 1, 9, 6, 8, 3, 7, 4, 2, 5\}$  при  $F = 300$ , приводит в совершенно другую область локального минимума. В другие области локального минимума приводит построение модифицированных времен в соответствии с эвристиками [6, 11] и др. (см. формулы (9)) и последующими шагами 6 алгоритма В. Например,  $\Pi = \{6, 2, 7, 1, 10, 8, 9, 3, 4, 5\}$  и  $F = 313$  при  $k = 1$ ;  $\Pi = \{10, 1, 9, 4, 3, 8, 6, 7, 5, 2\}$  и  $F = 320$  при  $k = 3$ ;  $\Pi = \{8, 6, 10, 1, 9, 7, 3, 5, 4, 2\}$  и  $F = 309$  при  $k = 2$ . Последнее расписание LV-алгоритмом преобразовывается в другое оптимальное решение:  $\Pi = \{8, 6, 10, 1, 9, 3, 7, 5, 4, 2\}$  и  $F = 289$ .

Таблица 2

## Эффективное и оптимальное решения примера

$i$	$T_1(i)$	$T_2(i)$	Последовательность	Времена завершения операций и заданий				Оптимальная последовательность	Времена завершения операций и заданий			
				$k = 1$	$k = 2$	$k = 3$	$k = 4$		$k = 1$	$k = 2$	$k = 3$	$k = 4$
1	194	246	6	10	30	60	70	6	10	30	60	70
2	194	176	10	25	48	70	115	10	25	48	70	115
3	288	277	8	43	60	100	130	8	43	60	100	130
4	203	167	1	58	78	125	160	1	58	78	125	160
5	205	95	9	78	103	155	185	9	78	103	155	185
6	170	180	3	103	133	195	213	3	103	133	195	213
7	255	245	7	118	168	235	245	7	118	168	235	245
8	183	192	2	130	195	265	270	4	153	173	249	269
9	240	260	4	165	200	279	299	2	165	200	279	284
10	179	261	5	205	215	284	304	5	205	215	284	289

Из двух различных оптимальных решений  $\Pi_1 = \{6, 10, 8, 1, 9, 3, 7, 4, 2, 5\}$  и  $\Pi_2 = \{8, 6, 10, 1, 9, 3, 7, 5, 4, 2\}$ , различающихся последовательностями выполнения заданий, может быть выбрано любое из них, обеспечивающее лучшее значение какого-то другого критерия эффективности.

Приведенный числовой пример наглядно иллюстрирует целесообразность использования в одном алгоритме нескольких различных эвристик с последующим локальным улучшением каждого из полученных решений.

## 5. РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Автором был проведен вычислительный эксперимент по решению тестовых задач большой размерности. Размерность задач — число заданий  $m$  и число машин  $n$  — выбиралась случайным образом. Во всех рассмотренных примерах ограничения на сроки завершения выполнения заданий и времена работы машин не учитывались. Времена выполнения различных операций в каждом из заданий выбирались случайным образом в диапазоне  $t_{ik} \in [5, 25]$ . Каждый из примеров был решен с применением различных эвристик, после чего полученное расписание преобразовывалось описанными алгоритмами локальной вариации (спуск в соответствующую область локального минимума). Результаты решения некоторых характерных примеров сведены в табл. 3. Жирным шрифтом выделены полученные наиболее эффективные приближенные решения, а курсивом — наилучшее из решений, полученных различными эвристическими алгоритмами.

Приведенная таблица показывает, что эффективность каждой из известных в литературе эвристик для различных задач различна, с помощью каждой из них могут быть получены отличающиеся друг от друга решения, принадлежащие различным областям притяжения локального минимума. Метод локальной вариации позволяет зачастую на 2—6 % повысить эффективность полученного эвристическими алгоритмами решения. Причем в некоторых случаях полученное даже существенно менее эффективное, чем на основе других эвристик, решение предложенными в работе методами локальной вариации может быть преобразовано в самое эффективное расписание. Учитывая, что значение критерия оптимальности в точном решении задачи, как правило, на 2—4 % превосходит значение нижней границы, вычисленной на начальном этапе решения, полученные результаты вычислительных экспериментов свидетельствуют об эффективности предлагаемых приближенных методов.

## ЗАКЛЮЧЕНИЕ

Разработаны модифицированные алгоритмы локальной вариации построенных последовательностей выполнения заданий, гарантирующие сохранение всех расписаний, содержащих допустимые и оптимальные решения, и осуществляющие спуск в точку локального минимума. Предложенные алгоритмы требуют существенно меньшего объема вычислений по сравнению с применяемыми на практике алгоритмами локального улучшения построенных расписаний.

Для традиционной flow-shop problem (при отсутствии ограничений на сроки выполнения зада-

Таблица 3

Результаты вычислительных экспериментов: эффективность применения различных эвристик и метода локальных вариаций

№ п/п	Размерность задачи		Отклонения (в процентах) полученного различными алгоритмами решения от нижней границы длины оптимального расписания (до и после применения алгоритмов локальной вариации)							
	$m$	$n$	[6]		[7]		[20]		[11]	
			до	после	до	после	до	после	до	после
1	25	6	7,3	<b>1,88</b>	8,1	3,57	<i>6,4</i>	2,42	9,35	3,26
2	50	7	6,43	4,37	7,12	4,18	7,81	5,23	<i>5,16</i>	<b>3,92</b>
3	70	8	8,56	<b>3,38</b>	<i>8,44</i>	5,66	10,13	4,76	9,64	6,28
4	85	5	<i>7,81</i>	6,43	8,71	5,21	9,05	<b>5,07</b>	10,12	7,32
5	100	7	<i>9,38</i>	6,98	9,67	<b>6,33</b>	11,24	8,43	10,83	7,52
6	110	10	11,19	7,05	12,32	<b>6,89</b>	<i>9,68</i>	7,59	11,82	8,17
7	120	6	<i>9,87</i>	7,28	11,43	8,14	10,77	<b>7,13</b>	9,89	7,93
8	135	5	<i>10,33</i>	9,15	11,01	8,42	10,81	<b>6,98</b>	11,14	7,76



ний и времени работы машин) разработаны комплексные алгоритмы решения, использующие основанные на различных гипотезах и известные в литературе эвристики и правила предпочтения и позволяющие осуществить попадание в различные области притяжения локальных минимумов. Эффективность работы алгоритмов повышается благодаря предложенным в работе методам локальной вариации и сравнения длины построенных расписаний.

Разработано программное обеспечение, позволяющее решать практические задачи размерностью до нескольких сотен заданий и выводить результаты решения как в форме таблиц, так и в графической форме — в виде диаграмм Ганта.

## ЛИТЕРАТУРА

1. *Зак Ю.А.* Решение обобщенной задачи Джонсона с ограничениями на сроки выполнения заданий и времена работы машин. Ч. 1. Точные методы решения // Проблемы управления. — 2010. — № 3. — С. 17—25.
2. *Johnson S.M.* Optimal two- and tree-stage production schedules with setup times included // Research Logistics Quarterly. — 1954 — Vol. 1. — P. 61—68.
3. *Domschke W., Scholl A., VoЯ S.* Produktionsplanung. Ablauforganisatorische Aspekte. — Berlin: Springer Verlag, 1997. — 456 p.
4. *Hundal T.S., Rajgopal J.* An extension of Palmer's heuristic for the flow-shop scheduling problem // International Journal of Production Research. — 1988. — Vol. 26. — P. 1119—1124.
5. *Gupta J.N.D.* A functional heuristic algorithm for the flow-shop scheduling problem // Operational Research Quarterly. — 1971. — Vol. 2. — P. 39—47.
6. *Cambell H.G., Dudek R.A., Smith M.L.* A heuristic algorithm for the  $n$  job,  $m$  machine sequencing problem // Management Science. — 1970. — Vol. 16. — P. 630—637.
7. *Dannenbring D.G.* A evaluation of flow shop sequencing heuristics // Management Science. — 1977. — Vol. 23. — P. 1174—1182.
8. *Ho J.C., Chang Y.-L.* A new heuristic for the  $n$ -job,  $M$ -machine problem // European Journal of Operational Research. — 1991. — Vol. 52. — P. 194—202.
9. *Ishibuchi H., Misaki S., Tanaka H.*: Modified simulated annealing algorithms for the flow shop sequencing problem // European Journal of Operational Research. — 1995. — Vol. 81. — P. 388—398.
10. *Ogbu F.A., Smith D.K.* Simulated annealing for the permutation flow shop problem // OMEGA. — 1991. — P. 64—67.
11. *Nawaz M., Enscore E.E., Ham I.* A heuristic algorithm for the  $m$ -machine,  $n$ -job flow-shop sequencing problem // OMEGA. — 1983. — Vol. 11. — P. 91—95.
12. *Brucker P.* Scheduling algorithms. — Berlin: Springer Verlag, 1995.
13. *Blazewicz J., Domschke W., Pesch E.* The job shop scheduling problem: Conventional and new solution techniques // European Journal of Operational Research. — 1996. — Vol. 93. — P. 1—33.
14. *Lageweg B.J., Lenstra J.K., Rinnooy Kann A.H.G.* A general bounding scheme for the permutation flow-shop problem // Operations Research. — 1978. — Vol. 26. — P. 53—67.
15. *Танаев В.С., Ковалёв М.Я., Шафранский Я.М.* Теория расписаний. Групповые технологии. — Минск: Ин-т техн. кибернетики НАН Беларуси, 1998. — 289 с.
16. *Танаев В.С., Сотсков Ю.Н., Струевич В.А.* Теория расписаний. Многостадийные системы. — М., URSS, 1989. — 328 с.
17. *Ogbu F.A., Smith D.K.* The application of the simulated annealing algorithm to the solution of the  $n/m/C_{\max}$  flow shop problem // Computer & Operations Research. — 1990. — Vol. 17. — P. 243—253.
18. *Cleveland G.A., Smith S.F.* Using genetic algorithms to schedule flow shop releases // Proc. of the third international conference on genetic algorithms, Morgan Kaufmann, San Mateo, 1989. — P. 160—169.
19. *Reeves C.R.* A genetic algorithm for flow shop sequencing // Computer & Operations Research. — 1995. — Vol. 22. — P. 5—13.
20. *Palmer D.S.* Sequencing job through a multi-stage process in the minimum total time — a quick method of obtaining a near optimum // Operations Research Quarterly. — 1965. — Vol. 16. — P. 101—107.

Статья представлена к публикации членом редколлегии В.Н. Бурковым.

**Зак Юрий Александрович** — д-р техн. наук, пенсионер, последнее место работы — Европейский центр по мехатронике (European Centre for Mechatronics) г. Аахен, Германия, ☎ +49 (0) 241/543255, ✉ yuriy\_zack@hotmail.com.

### НОВАЯ КНИГА

**Левин В.И. Логические методы в теории надежности: монография.** — Пенза: Изд-во Пензенской гос. технол. академии, 2010. — 110 с.

Разработана математическая модель надежности сложных систем управления, пригодная для создания эффективных и экономичных методов количественного изучения надежности таких систем. Она представляет собой динамико-автоматную модель, описывающую надежностные процессы в системах в виде функций от надежностных процессов в элементах этих систем. Разработан математический аппарат непрерывной логики, позволяющий адекватно описывать созданную динамико-автоматную модель надежностных процессов в системах управления. Благодаря этому стали возможными практическое моделирование, расчет и анализ надежности указанных систем. Разработанный аппарат представляет собой обобщение булевой логики на непрерывный случай.

Книгу можно заказать по e-mail: [f\\_os@pgta.ru](mailto:f_os@pgta.ru)

# ИНВЕСТИЦИИ И СБАЛАНСИРОВАННЫЙ РОСТ В МОДЕЛИ ДЕЦЕНТРАЛИЗОВАННОЙ ЭКОНОМИКИ<sup>1</sup>

А.П. Абрамов

Рассмотрены две схемы выхода на неймановскую траекторию в модели децентрализованной экономики с учетом инвестиций. Первая из них базируется на натуральных показателях, вторая — на денежных расчетах хозяйствующих субъектов. Приведены условия, при которых эти схемы обеспечивают асимптотическое решение указанной задачи.

**Ключевые слова:** децентрализованная экономика, сбалансированный рост, инвестиции.

## ВВЕДЕНИЕ

Теория сбалансированного роста — классический раздел математической экономики, с ее основными результатами можно ознакомиться в работе [1]. Модели экономики, изучаемые этой теорией, неявно предполагают существование единого управляющего центра, обладающего всей информацией об экономической системе. Этот центр полностью контролирует производство и распределение продукции, а все его решения выполняются абсолютно точно. Ясно, что подобные модели не соответствуют принципам рыночной экономики, в которой хозяйствующие субъекты самостоятельно планируют и организуют свою работу. Один из возможных подходов к построению общей схемы линейной теории экономического роста для децентрализованной экономики предложен в статье [2]. Он базируется на стандартной гейловской теории в совокупности с моделью равновесия Вальраса. Однако анализ данного подхода показывает необходимость централизованного распределения избыточных продуктов. Альтернативный подход, не использующий равновесие по Вальрасу, предложен в работе [3]. В настоящей статье развивается это направление и исследуются возможные схемы выхода на неймановскую траекторию децентрализованной экономики с учетом фактора инвестиций.

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проект 10-07-00286).

## 1. ОПИСАНИЕ МОДЕЛИ

Рассмотрим замкнутую динамическую модель производства и товарообмена, в которой фигурирует  $n$  монопродуктовых отраслей, причем каждый из видов продукции производится только одной отраслью. Состояние данной экономической системы отслеживается в дискретные моменты времени, которые обозначаются индексом  $t$ ,  $t = 0, 1, 2, \dots$ . Шаги модели, т. е. промежутки времени между соседними моментами, будем помечать тем же индексом  $t$ , причем номер шага соответствует правой границе. Длительности всех шагов предполагаются одинаковыми и равными одному производственному циклу во всех отраслях. Продукция, произведенная на некотором шаге, должна быть использована до окончания следующего шага.

Введем обозначения:

$i = 1, \dots, n$  — индекс отрасли;

$N_i^{r+}$  — подмножество отраслей, продукция которых необходима отрасли  $i$  в качестве сырья или комплектующих для выпуска продукции;

$N_i^{k+}$  — подмножество отраслей, продукция которых необходима отрасли  $i$  для наращивания производственных мощностей;

$N_i^{r-}$  — подмножество отраслей, потребляющих продукцию отрасли  $i$  в качестве сырья или комплектующих;

$N_i^{k-}$  — подмножество отраслей, потребляющих продукцию отрасли  $i$  в качестве фондообразующей;



$\xi_i(t)$  — производственные мощности отрасли  $i$  на шаге  $t$ , равные максимально возможному объему производства на данном шаге;

$x_i(t)$  — объем выпуска продукции отраслью  $i$  на шаге  $t$ ;

$y_{ji}(t)$  — объем производственного ресурса вида  $j$ ,  $j \in N_i^{r+}$ , которым располагает отрасль  $i$  в начале шага  $t$ ;

$z_{ji}(t)$  — объем капитального ресурса вида  $j$ ,  $j \in N_i^{k+}$ , которым располагает отрасль  $i$  в начале шага  $t$ .

Вообще говоря, номенклатуры производственных и капитальных ресурсов не совпадают. Однако чтобы не усложнять описание модели, явно этот факт не будет учитываться. Предполагается, что множества  $N_i^{r+}$ ,  $N_i^{k+}$ ,  $N_i^{r-}$ ,  $N_i^{k-}$ ,  $i = 1, \dots, n$  не пусты, а производственная функция отрасли  $i$  имеет вид

$$x_i(t) = \min \left\{ \xi_i(t), \min_{j \in N_i^{r+}} \left\{ \frac{y_{ji}(t)}{y_{ji}^0} \right\} \right\}, \quad i = 1, \dots, n, \quad (1)$$

где  $y_{ji}^0 > 0$  — минимально необходимое количество продукта вида  $j$ , которое требуется для производства одной единицы продукции вида  $i$ . Эту функцию называют производственной функцией Леонтьева, а также производственной функцией с фиксированными пропорциями факторов [4, 5].

Выпишем балансовое уравнение для продукции, произведенной отраслью  $i$  на шаге  $t$ :

$$x_i(t) = \sum_{j \in N_i^{r-}} y_{ij}(t+1) + \sum_{j \in N_i^{k-}} z_{ij}(t+1) + v_i(t), \quad (2)$$

где переменная  $v_i(t)$ , равна объему нераспределенной продукции.

Динамику производственных мощностей отрасли  $i$  будем описывать уравнением вида

$$\xi_i(t) = (1 - \mu)\xi_i(t-1) + \min_{j \in N_i^{k+}} \left\{ \frac{z_{ji}(t-\tau)}{z_{ji}^0} \right\}, \quad i = 1, \dots, n, \quad (3)$$

где  $\mu$  — коэффициент амортизации, который считается одинаковым для всех отраслей;  $z_{ji}^0$  — минимально необходимый объем продукции вида  $j$ ; необходимый для увеличения на единицу производственных мощностей отрасли  $i$ ,  $\tau$  — временной лаг, характеризующий запаздывание в создании новых мощностей. Этот параметр предполагается одинаковым для всех отраслей и всех видов фондообразующей продукции и равным числу шагов, за которое происходит освоение этой продукции. Случай  $\tau = 0$  означает возможность ее использования

в качестве основных производственных фондов уже на следующем производственном цикле. Предположим, что система функционирует так, что на всех шагах вся продукция полностью распределяется, и в процессах (1) и (3) нет избыточных ресурсов (в том числе мощностей в процессе (1)). Это означает, что  $x_i(t) = \xi_i(t)$  при всех  $i$  и  $t$ . В этом случае из выражений (1)–(3) имеем матричное уравнение вида

$$x(t) = Yx(t+1) + Z(x(t+\tau+1) - (1-\mu)x(t+\tau)), \quad (4)$$

где  $x(\cdot)$  — вектор-столбцы, а элементы  $\bar{y}_{ij}$  и  $\bar{z}_{ij}$  квадратных матриц  $Y$  и  $Z$  порядка  $n$  определены соответственно так:

$$\bar{y}_{ij} = \begin{cases} y_{ij}^0, & i \in N_j^{r+}; \\ 0, & i \notin N_j^{r+}, \end{cases} \quad \bar{z}_{ij} = \begin{cases} z_{ij}^0, & i \in N_j^{k+}; \\ 0, & i \notin N_j^{k+}. \end{cases}$$

Легко видеть, что указанный режим может быть обеспечен при постоянном темпе роста  $\gamma$  объемов выпуска всех видов продукции:  $x_i(t+1)/x_i(t) = \gamma$  для всех  $i$  и  $t$ . В этом случае уравнение (4) принимает вид

$$(1/\gamma)x(t) = (Y + (\gamma^\tau - (1-\mu)\gamma^{\tau-1})Z)x(t).$$

Оно показывает, что достаточное условие такого режима состоит в наличии у матрицы

$$M(\gamma) = Y + (\gamma^\tau - (1-\mu)\gamma^{\tau-1})Z \quad (5)$$

положительного собственного числа  $1/\gamma$  и соответствующего строго положительного собственного вектора  $x$ , а начальный вектор  $x(0)$  должен быть коллинеарен вектору  $x$ . Поскольку интерес представляют только экономические системы с возможностью расширенного воспроизводства, сформулируем условия, при которых существует такое  $\gamma > 1$ , что матрица  $M(\gamma)$  имеет собственное число  $1/\gamma$ . Для удобства заменим в матрице (5) параметр  $\gamma$  на параметр  $\lambda = 1/\gamma$  и рассмотрим матрицу  $\tilde{M}(\lambda) \equiv M(1/\lambda)$ .

**Теорема 1.** Пусть матрица  $(Y+Z)$  неразложима, и все ее столбцевые суммы строго меньше единицы, т. е.

$$\sum_{i=1}^n (\bar{y}_{ij} + \bar{z}_{ij}) < 1, \quad j = 1, \dots, n.$$

Тогда существует скаляр  $\lambda^* \in (0, 1)$ , являющийся собственным числом матрицы  $\tilde{M}(\lambda^*)$ , которому отвечает единственный (с точностью до скалярного множителя) собственный вектор  $x^*$ , все координаты которого ненулевые и одного знака. Если найдется такой скаляр  $\tilde{\lambda} \in (0, \lambda^*)$ , то он также является собственным числом матрицы  $\tilde{M}(\tilde{\lambda})$ , то любой со-

ответствующий ему собственный вектор имеет компоненты разных знаков. ♦

Доказательство этой теоремы аналогично доказательству теоремы 2.6 в книге [6, гл. 2]. Таким образом, если  $x(0)$  — фробениусов вектор матрицы  $\tilde{M}(\lambda^*)$ , то экономическая система может функционировать в режиме *неймановского процесса*, т. е. иметь *сбалансированный рост с максимальным темпом*  $\gamma$  [6]. При этом траектория выпусков будет принадлежать *неймановскому лучу*, называемому также *магистралью* [1]. Для краткости, указанный режим функционирования будем именовать *магистральным*.

Всюду далее предполагается, что параметры экономической системы удовлетворяют условиям теоремы 1 и, значит,  $\gamma \equiv 1/\lambda^* > 1$ .

## 2. ПЛАНИРОВАНИЕ НА ОСНОВЕ МАКСИМИЗАЦИИ ПРИБЫЛИ

До сих пор в модели неявно предполагалось наличие централизованной системы управления, которая полностью контролирует производство и распределение продукции. Далее будем считать, что отрасли работают в условиях хозяйственной автономии. При этом каждая отрасль  $i$  стремится максимизировать свою прибыль  $\Pi_i(t)$  на текущем шаге  $t$ , определяемую как разность между выручкой и соответствующими затратами:

$$\begin{aligned} \Pi_i(t) = & p_i(t)x_i(t) - \sum_{j \in N_i^{r+}} p_j(t-1)y_{ji}(t) - \\ & - \sum_{j \in N_i^{k+}} p_j(t-1)z_{ji}(t), \end{aligned} \quad (6)$$

где цена  $p_i(t)$  связана с объемом продаж, равным на магистрали объему производства,  $x_i(t)$  линейным уравнением вида

$$\begin{aligned} p_i(t) = & a_i(t) + b_i(t)x_i(t), \\ i = & 1, \dots, n, \quad t = 0, 1, \dots, \end{aligned} \quad (7)$$

в котором параметры  $a_i(t) > 0$ ,  $b_i(t) < 0$  имеют фиксированные значения при данных  $i$  и  $t$ . Таким образом, отрицательное значение  $b_i(t)$  в уравнении (7) снижает цену с ростом объема продаж.

Ясно, что параметр  $a_i(t)$  равен предельному значению цены, когда объем реализации (производства) данного вида продукции стремится к нулю. Параметр  $b_i(t)$  указывает, насколько снижается цена при увеличении объема реализации (производства) на единицу. Будем называть их соответственно *базовой ценой* и *параметром скидки*. При этом цена должна быть положительной:

$$a_i(t) + b_i(t)x_i(t) > 0, \quad i = 1, \dots, n; \quad t = 0, 1, \dots$$

В магистральном режиме выражение (6) принимает вид

$$\begin{aligned} \Pi_i(t) = & (a_i(t) + b_i(t)x_i(t))x_i(t) - \sum_{j \in N_i^{r+}} (a_j(t-1) + \\ & + b_j(t-1)x_j(t-1))y_{ji}^0 x_i(t) - \gamma^{\tau-1}(\gamma - (1 - \mu)) \times \\ & \times \sum_{j \in N_i^{k+}} (a_j(t-1) + b_j(t-1)x_j(t-1))z_{ji}^0 x_i(t). \end{aligned}$$

Так как выбор переменной  $x_i(t)$  происходит при фиксированных ценах предыдущего шага, и  $b_i(t) < 0$ , то максимум прибыли отрасли  $i$  на шаге  $t$  достигается при

$$\begin{aligned} x_i(t) = & -\frac{1}{2b_i(t)} \left( a_i(t) - \sum_{j \in N_i^{r+}} (a_j(t-1) + \right. \\ & + b_j(t-1)x_j(t-1))y_{ji}^0 - \gamma^{\tau-1}(\gamma - (1 - \mu)) \times \\ & \left. \times \sum_{j \in N_i^{k+}} (a_j(t-1) + b_j(t-1)x_j(t-1))z_{ji}^0 \right). \end{aligned} \quad (8)$$

Отсюда следует, что в магистральном режиме коэффициенты  $a_i(t)$  и  $b_i(t)$  должны удовлетворять при всех  $i$  и  $t$  системе уравнений вида

$$\begin{aligned} a_i(t) + 2\gamma^t b_i(t)x_i(0) - \sum_{j \in N_i^{r+}} (a_j(t-1) + \\ + \gamma^{t-1} b_j(t-1)x_j(0))y_{ji}^0 - \gamma^{\tau-1}(\gamma - (1 - \mu)) \times \\ \times \sum_{j \in N_i^{k+}} (a_j(t-1) + \gamma^{t-1} b_j(t-1)x_j(0))z_{ji}^0 = 0. \end{aligned}$$

Для этого достаточно, чтобы переменные  $a_i(t)$  удовлетворяли системе уравнений

$$\begin{aligned} a_i(t) = & \sum_{j \in N_i^{r+}} y_{ji}^0 a_j(t-1) + \gamma^{\tau-1}(\gamma - (1 - \mu)) \times \\ & \times \sum_{j \in N_i^{k+}} z_{ji}^0 a_j(t-1), \end{aligned} \quad (9)$$

а переменные  $b_i(t)$  — системе уравнений

$$\begin{aligned} b_i(t) = & \frac{1}{2\gamma x_i(0)} \left( \sum_{j \in N_i^{r+}} y_{ji}^0 x_j(0) b_j(t-1) + \right. \\ & \left. + \gamma^{\tau-1}(\gamma - (1 - \mu)) \sum_{j \in N_i^{k+}} z_{ji}^0 x_j(0) b_j(t-1) \right). \end{aligned} \quad (10)$$

Обозначим через  $a(0)$  и  $b(0)$  вектор-строки, образованные из начальных значений ценовых коэффициентов. Пусть  $a(0) = p^*$ , где  $p^*$  — левый фробениусов вектор матрицы  $\tilde{M}(\lambda^*)$ , а  $b(0)$  — левый фробениусов вектор матрицы  $\hat{M}(\lambda^*)$ , элементы



которой связаны с элементами матрицы  $\tilde{M}(\lambda^*)$  равенствами вида  $\hat{m}_{ij} = x_i(0)/x_j(0)\tilde{m}_{ij}$ . При  $t \geq 1$  положим

$$a_i(t) = \gamma^{-1}a_i(t-1), \quad b_i(t) = 0,5\gamma^{-2}b_i(t-1), \\ i = 1, \dots, n. \quad (11)$$

Элементарные вычисления показывают, что при такой динамике ценовых коэффициентов в магистральном режиме справедливы следующие утверждения:

- выполняются динамические уравнения (9) и (10);
- если  $a_i(0)$  и  $b_i(0)$  таковы, что  $p_i(0) > 0$ , то  $p_i(t) > 0$  при всех  $t \geq 1$ ;
- все отрасли имеют положительную прибыль на всех шагах  $\Pi_i(t) = -b_i(t)x_i^2(t) > 0$ ;
- все отрасли имеют нулевые платежные балансы  $B_i(t)$  на всех шагах

$$B_i(t) \equiv (a_i(t) + b_i(t)x_i(t))x_i(t) - \\ - \sum_{j \in N_i^{r+}} (a_j(t) + b_j(t)x_j(t))y_{ji}(t+1) - \\ - \sum_{j \in N_i^{k+}} (a_j(t) + b_j(t)x_j(t))z_{ji}(t+1) = 0.$$

Отметим, что положительная прибыль всех отраслей на каждом шаге не противоречит последнему равенству, так как закупка ресурсов и продажа произведенной продукции происходят со сдвигом в один временной шаг.

Легко видеть, что планирование выпуска согласно выражению (8) в совокупности с соотношениями (11) обеспечивает системе магистральное режим при соответствующих начальных условиях.

### 3. ПЕРВАЯ СХЕМА ВЫХОДА НА НЕЙМАНОВСКУЮ ТРАЕКТОРИЮ

Рассмотренный в § 2 подход решает задачу удержания экономической системы на луче Неймана. Ясно, что для его применения векторы  $x(0)$ ,  $\xi(1)$ , ...,  $\xi(\tau-2)$  должны иметь согласованные значения, а всем отраслям должен быть известен максимальный темп сбалансированного роста  $\gamma$ . Если требуется, чтобы все отрасли имели нулевые платежные балансы на всех шагах, то вектор-строки  $a(0)$  и  $-b(0)$ , образованные из соответствующих ценовых коэффициентов, должны быть фробениусовыми векторами для матриц  $\tilde{M}(\lambda^*)$  и  $\hat{M}(\lambda^*)$  соответственно. В этом случае отрасли, определяя цены на свою продукцию на основе динамических уравнений (11), и планируя производство исходя из максимума прибыли, обеспечивают системе магистральное режим.

Выполнение на практике всех этих условий представляется маловероятным. Кроме того, любые сбои в работе системы сразу уведут систему с магистрали, при этом для некоторых видов продукции возникает превышение спроса над предложением. В этих обстоятельствах данный механизм планирования перестает работать ввиду отсутствия процедур принятия управленческих решений в таких ситуациях. Поэтому заслуживают изучения схемы планирования со следующими свойствами:

- не требуют централизованного управления;
- могут вырабатывать решения вне магистрального режима и при этом (асимптотически) выводят систему на магистраль;
- не уведут систему с магистрали, если она работает в этом режиме;
- имеют ясное экономическое обоснование.

Описание работы многосекторной экономики вне магистрального режима требует введения дополнительных переменных для отрасли  $i$ ,  $i = 1, \dots, n$ , в которых явно проявляется ее экономическая самостоятельность, и которые должны быть определены к моменту перехода от шага  $t-1$  к шагу  $t$ :

$x_i^p(t)$  — планируемый объем выпуска на шаге  $t$ ;

$\Delta k_i^p(t+\tau)$  — план ввода производственных мощностей к началу шага  $t+\tau$ ;

$x_i^s(t-2)$  — объем реализованной (отгруженной потребителям) продукции, которая была произведена на шаге  $t-2$ .

Совокупность указанных планов позволяет определить для отрасли  $i$  показатель  $x_i^d(t-1)$  — суммарный спрос потребителей на продукцию шага  $t-1$ , вычисляемый по формуле

$$x_i^d(t-1) = \sum_{j \in N_i^{r-}} y_{ij}^0 x_j^p(t) + \sum_{j \in N_i^{k-}} z_{ij}^0 \Delta k_j^p(t+\tau), \\ i = 1, \dots, n,$$

которая предполагает, что все потребители стремятся достичь планируемых показателей наиболее экономным способом.

Опишем схему функционирования данной экономической системы, которая при некоторых условиях выводит систему на магистраль, а также обеспечивает сохранение магистрального режима.

**Базовая схема функционирования системы.** Первоначальный план выпуска отрасли  $i$  на шаге  $t$ , а также план ввода мощностей этой отрасли к началу шага  $t+\tau$  однозначно определяются объемом реализации продукции, произведенной на шаге  $t-2$ :

$$x_i^p(t) = \alpha x_i^s(t-2), \quad \Delta k_i^p(t+\tau) = \beta x_i^s(t-2), \\ t \geq 2, \quad i = 1, \dots, n, \quad (12)$$

где коэффициенты  $\alpha$  и  $\beta$  одни и те же для всех отраслей на всех шагах. Наряду с показателями (12) отрасль  $i$  вычисляет коэффициент

$$\chi_i(t) = x_i^p(t)/\xi_i(t), \quad i = 1, \dots, n,$$

который измеряет первоначальный план выпуска с имеющимися мощностями. Далее, суммируя заявки от потребителей, отрасль  $i$  определяет показатель  $x_i^d(t-1)$  первоначального спроса на свою продукцию. Поскольку в момент перехода к шагу  $t$  уже известен реальный объем выпуска  $x_i^p(t-1)$ , то отрасль  $t$  вычисляет коэффициент

$$\eta_i(t) = x_i^d(t-1)/x_i^p(t-1), \quad i = 1, \dots, n,$$

характеризующий обеспеченность планов ресурсом, который она произвела. Далее все отрасли обмениваются коэффициентами  $\chi_i(t)$ ,  $\eta_i(t)$  и каждая из отраслей (или некий информационный центр) вычисляет максимальное значение этих показателей:

$$\sigma(t) = \max_i \{\chi_i(t), \eta_i(t)\}.$$

Если окажется, что  $\sigma(t) \leq 1$ , то производственные мощности позволяют выполнить намеченные планы, спрос каждой из отраслей на ресурсы удовлетворяется полностью, объемы поставок определяются так:

$$y_{ij}(t) = y_{ij}^0 x_j^p(t), \quad j \in N_i^{r-},$$

$$z_{ij}(t) = z_{ij}^0 \Delta k_j^p(t + \tau), \quad j \in N_i^{k-}, \quad i = 1, \dots, n,$$

и выпуск продукции равен первоначальному плану:  $x_i^p(t) = x_i^p(t)$ ,  $i = 1, \dots, n$ .

Ясно, что при  $\sigma(t) > 1$  полное выполнение исходных планов становится невозможным. В этом случае все отрасли, используя данный параметр, уменьшают планы выпусков и планы развития:

$$x_i^p(t) := x_i^p(t)/\sigma(t), \quad \Delta k_i^p(t + \tau) := \Delta k_i^p(t + \tau)/\sigma(t), \quad i = 1, \dots, n. \quad (13)$$

Соответственно, потребность в мощностях и спрос на все ресурсы также уменьшаются в  $\sigma(t)$  раз, все скорректированные планы полностью выполнимы, а вновь вычисленный показатель  $\sigma(t)$  равен единице.

Если причина корректировки состоит в нехватке ресурсов, то выражение (13) означает, что планы пересчитываются по наиболее дефицитному ресурсу, и этот ресурс распределяется пропорционально размерам спроса. Тем самым в модели ни у одной из отраслей нет привилегий по ресурсному обеспечению.

План, допустимый по ресурсам, однозначно определяет для отрасли  $i$  показатель  $x_i^s(t-1)$  —

объем реализации продукции, произведенной на шаге  $t-1$ :

$$x_i^s(t-1) = \sum_{j \in N_i^{r-}} y_{ij}^0 x_j^p(t) + \sum_{j \in N_i^{k-}} z_{ij}^0 \Delta k_j^p(t + \tau), \quad i = 1, \dots, n.$$

Далее начинается производственный цикл на шаге  $t$ , после завершения которого по показателям (12) определяются начальные значения векторов  $x^p(t+1)$ ,  $\Delta k_i^p(t + \tau + 1)$  и т. д. Что касается начала работы схемы при  $t=2$ , то вместо  $x_i^s(0)$  в выражении (12) можно выбрать параметр  $\xi_i(0)$ . Кроме того, должны быть заданы векторы  $x(1)$ ,  $\xi(2)$ , ...,  $\xi(\tau + 1)$ . Легко видеть, что базовая схема функционирования системы позволяет вырабатывать все необходимые управленческие решения вне магистрали. С другой стороны, элементарная индукция показывает, что экономическая система будет функционировать в магистральном режиме, если выполнены следующие условия:

$$\beta/\alpha = \gamma^{\tau-1}(\gamma - (1 - \mu)), \quad \alpha \geq \gamma^2, \quad (14)$$

а) векторы  $x(1) \equiv \xi(1) \equiv \gamma \xi(0)$  суть фробениусовы матрицы  $M(\gamma)$ ,

б) производственные мощности растут с темпом  $\gamma$  на шагах  $2, \dots, \tau + 1$ .

В общем случае имеет место следующий результат.

**Теорема 2.** Пусть в условиях теоремы 1 матрица  $Y + Z$  примитивна, векторы  $x(1)$ ,  $\xi(1)$ , ...,  $\xi(\tau + 1)$  строго положительны и выполняются ограничения (14). Тогда базовая схема функционирования либо асимптотически выводит систему на магистраль, либо оставляет систему на магистрали при выполнении условий а) и б). ♦

**Доказательство.** Введем параметр  $\rho(t)$ , показывающий степень выполнения первоначальных планов на шаге  $t$ :

$$\rho(t) = \begin{cases} 1, & \sigma(t) \leq 1; \\ 1/\sigma(t), & \sigma(t) > 1. \end{cases}$$

Применяя индукцию, выпишем зависимость между векторами  $x(t)$  и  $x^p(2)$ :

$$x(t) = \rho(t)\rho(t-1)\dots\rho(2)\alpha^{t-2} \tilde{M}(\lambda^*)^{t-2} x^p(2).$$

Используя фробениусово число  $\lambda^*$  матрицы  $\tilde{M}(\lambda^*)$ , запишем последнее равенство так:

$$x(t) = \rho(t)\rho(t-1)\dots\rho(2)(\lambda^*)^{t-2} \left( \frac{\tilde{M}(\lambda^*)}{\lambda^*} \right)^{t-2} x^p(2).$$

Так как матрица  $\tilde{M}(\lambda^*)$  неразложима и примитивна, то (см. книгу [6]) последовательность



$(\tilde{M}(\lambda^*)/\lambda^*)^t x^p(2)$  сходится к пределу  $\nu x^*$  при  $t \rightarrow \infty$ , где  $\nu = \|x^p(2)\|/\|x^*\|$ , а  $x^*$  — фробениусов вектор матрицы  $M(\lambda^*)$ . Это означает, что последовательность нормированных векторов вида  $x(t)/\|x(t)\|$  при  $t \rightarrow \infty$  сходится к пределу  $x^*/\|x^*\|$ , т. е. к нормированному фробениусову вектору матрицы  $\tilde{M}(\lambda^*)$ . Отсюда следует, что равенство  $x(t+1) = \gamma x(t)$  будет выполняться сколь угодно точно при достаточно больших  $t$ . ♦

#### 4. ВТОРАЯ СХЕМА ВЫХОДА НА НЕЙМАНОВСКУЮ ТРАЕКТОРИЮ

В базовой схеме функционирования экономической системы применялись только натуральные показатели. Рассмотрим схему выхода на магистраль, основанную на финансовых показателях. Предположим, что в каждой отрасли есть собственник, которому принадлежат производственные мощности, а также (наемный) управляющий, который занимается текущей работой данной отрасли. Последний берет в аренду производственные мощности у собственника, выплачивая ему определенные платежи. В свою очередь, владелец капитала контролирует динамику мощностей посредством закупок фондообразующей продукции.

Конкретно, будем предполагать, что в момент перехода от шага  $t-1$  к шагу  $t$  управляющий должен определить показатель  $\Pi_i^{rp}(t)$  — планируемую величину производственной прибыли отрасли  $i$  на шаге  $t$ . Оценивая показатель  $\Pi_i^{rp}(t)$ , он полагает, что будет продан весь объем продукции, планируемой к выпуску. Кроме того, ему известна функция  $\pi_i(t)$  — размер платежа, которую установил владелец капитала за использование единицы производственных мощностей отрасли  $i$  на шаге  $t$ . Поскольку использование капитальных и производственных ресурсов сверх необходимого минимума приводит к денежным потерям, то данный показатель вычисляется так:

$$\Pi_i^{rp}(t) = p_i(t)x_i^p(t) - \sum_{j \in N_i^{r+}} p_j(t-1)y_{ji}^0 x_i^p(t) - \pi_i(t)x_i^p(t). \quad (15)$$

Предположим, что цены  $p_i(t)$ ,  $\pi_i(t)$  отвечают принципу «базовая ставка — скидка» и являются линейными функциями вида

$$\begin{aligned} p_i(t) &= a_i^r(t) + b_i^r(t)x_i(t), \\ \pi_i(t) &= a_i^k(t) + b_i^k(t)\xi_i(t), \end{aligned} \quad (16)$$

где  $a_i^r(t) > 0$ ,  $a_i^k(t) > 0$ ,  $b_i^r(t) < 0$ ,  $b_i^k(t) < 0$ , а  $\xi_i(t)$  — объем производственных мощностей, сданных в аренду. Как и ранее, функции  $p_i(t)$ ,  $\pi_i(t)$  могут принимать только положительные значения. Если  $b_i^r(t) < b_i^k(t)$ , то показатель (15) достигает максимума при

$$x_i^p(t) = \frac{1}{2(b_i^r(t) - b_i^k(t))} \times \left( a_i^r(t) - \sum_{j \in N_i^{r+}} p_j(t-1)y_{ji}^0 - a_i^k(t) \right). \quad (17)$$

Предположим, что параметр  $b_i^r(t)$  задан. Определим параметр  $a_i^r(t)$  так:

$$\begin{aligned} a_i^r(t) &= \sum_{j \in N_i^{r+}} p_j(t-1)y_{ji}^0 + a_i^k(t) - \\ &- 2(b_i^r(t) - b_i^k(t))\alpha x_i^s(t-2), \end{aligned} \quad (18)$$

где коэффициент  $\alpha$  тот же самый, что и в показателях (12). Подставляя это значение в формулу (17), получаем значение  $x_i^p(t)$  из выражения (12). При этом показатель (15) принимает вид

$$\Pi_i^{rp}(t) = -(b_i^r(t) - b_i^k(t))(x_i^p(t))^2, \quad (19)$$

и планируемая прибыль будет положительной, если  $|b_i^r(t)| > |b_i^k(t)|$ .

Что касается владельца капитала отрасли  $i$ , то в начале шага  $t$  он планирует финансовые траты в объеме  $\sum_{j \in N_i^{k+}} p_j(t-1)z_{ji}^p(t) = \sum_{j \in N_i^{k+}} p_j(t-1)z_{ji}^0 \Delta k_i^p(t+\tau)$  на закупку фондообразующей продукции. Эти траты должны принести доход на шаге  $t+\tau$ , а также на всех последующих шагах. С другой стороны, доход от сдачи в аренду мощностей на шаге  $t+\tau$  зависит и от объема мощностей  $\xi_i(t+\tau-1)$ , которые образовались в результате предыдущих капиталовложений. Поэтому, для простоты, будем оценивать ожидаемый эффект от рассматриваемых затрат только на шаге  $t+\tau$ . При этом введем фактор дисконтирования  $\varepsilon = \varepsilon(\tau) < 1$ , который приводит ценность денег шага  $t+\tau$  к шагу  $t$ .

Представим планируемый объем мощностей  $\xi_i^p(t+\tau)$  отрасли  $i$  к началу шага  $t+\tau$  в виде

$$\xi_i^p(t+\tau) = (1-\mu)\xi_i^p(t+\tau-1) + \Delta k_i^p(t+\tau). \quad (20)$$

Собственник полагает, что этот объем будет равен соответствующему спросу на мощности.

Тогда планируемый размер прибыли  $\Pi_i^{rp}(t)$  оценивается так:

$$\Pi_i^{rp}(t) = \varepsilon \pi_i(t + \tau) \xi_i^p(t + \tau) - \sum_{j \in N_i^{k+}} p_j(t - 1) z_{ji}^p(t). \quad (21)$$

Из выражений (16), (20) и (21) следует, что максимум прибыли, как функции  $\Delta k_i^p(t + \tau)$ , достигается при

$$\Delta k_i^p(t + \tau) = \frac{a_i^k(t + \tau) - \varepsilon^{-1} \sum_{j \in N_i^{k+}} p_j(t - 1) z_{ji}^0 + 2b_i^k(t + \tau)(1 - \mu) \xi_i^p(t + \tau - 1)}{2b_i^k(t + \tau)}. \quad (22)$$

Предположим, что параметр  $b_i^k(t + \tau)$  задан.

Определим параметр  $a_i^k(t + \tau)$  так:

$$a_i^k(t + \tau) = \varepsilon^{-1} \sum_{j \in N_i^{k+}} p_j(t - 1) z_{ji}^0 - 2b_i^k(t + \tau) \times ((1 - \mu) \xi_i^p(t + \tau - 1) + \beta x_i^s(t - 2)), \quad (23)$$

где коэффициент  $\beta$  тот же самый, что и в выражении (12). Подстановка этого значения в формулу (22) дает  $\Delta k_i^p(t + \tau)$  из выражения (12). При этом формула (21) принимает вид

$$\Pi_i^{rp}(t) = -\varepsilon b_i^k(t + \tau) (\Delta k_i^p(t + \tau))^2 + \varepsilon (a_i^k(t + \tau) + b_i^k(t + \tau)(1 - \mu) \xi_i^p(t + \tau - 1)) (1 - \mu) \xi_i^p(t + \tau - 1).$$

Таким образом, в натуральных показателях данная схема идентична базовой схеме и, значит, также обеспечивает асимптотический выход на магистраль либо сохранение магистрального режима. Рассмотрим экономическую интерпретацию соотношений (18) и (23), сформулировав их в виде правил.

**Правило назначения базовой цены.** В базовую цену закладываются издержки на производство единицы продукции плюс базовая ставка арендной платы плюс компенсация в двойном размере ожидаемых удельных потерь из-за скидок при планируемом объеме продаж.

**Правило назначения базовой ставки арендной платы.** Базовая ставка арендной платы должна быть равна издержкам (с учетом дисконта) на ввод

в строй единицы мощностей плюс компенсация в двойном размере ожидаемых удельных потерь из-за скидок при сдаче мощностей в аренду.

Рассмотренный финансовый механизм допускает определенный произвол в выборе параметров  $b_i^r(t)$  и  $b_i^k(t)$ . Например, если их динамика описывается геометрической прогрессией вида  $b_i^m(t) = \gamma^{-t} b_i^m(0)$ ,  $m = r, k$ ;  $b_i^r(0) < b_i^k(0) < 0$ ,  $i = 1, \dots, n$ , то из соотношения (19) видно, что производственная прибыль каждой из отраслей асимптотически выходит на темп роста  $\gamma$ .

## ЗАКЛЮЧЕНИЕ

Рассмотренная модель показывает теоретическую возможность асимптотического выхода многосекторной экономики на неймановскую траекторию, когда отсутствует централизованное планирование и управление, а экономические агенты ориентируются на максимизацию своей прибыли. Тем не менее, некоторая координация планов работ отраслей необходима в рамках предложенной схемы. Задача координирующего органа состоит в определении «правильных» значений параметров  $\alpha$  и  $\beta$  (см. условия (14)) и передаче этих данных всем отраслям. Кроме того, координирующий орган должен установить такой порядок в сфере экономики, при котором ни одна из отраслей не имеет привилегий по ресурсному обеспечению.

## ЛИТЕРАТУРА

1. Никайдо Х. Выпуклые структуры и математическая экономика. — М.: Мир, 1972. — 520 с.
2. Бельский В.З., Слаников А.Д. Равновесная динамика замкнутого рынка монопродуктовых производств // Экономика и математические методы. — 1994. — № 4. — С. 112—128.
3. Абрамов А.П. О выходе на магистраль сбалансированного роста в модели замкнутой децентрализованной экономики // Математическое моделирование. — 2008. — № 2. — С. 3—12.
4. Иванилов Ю.П., Лотов А.В. Математические модели в экономике. — М.: Наука, 1979. — 304 с.
5. Клейнер Г.Б. Производственные функции: теория, методы, применение. — М.: Финансы и статистика, 1986. — 240 с.
6. Ашманов С.А. Введение в математическую экономику. — М.: Наука, 1984. — 294 с.

Статья представлена к публикации членом редколлегии В.Н. Бурковым.

Абрамов Александр Петрович — д-р физ.-мат. наук, гл. науч. сотрудник, Учреждение Российской академии наук Вычислительный центр им. А.А. Дородницына РАН, ☎ (499) 135-00-80, ✉ apabramov@list.ru, apabra@ccas.ru.



# ПРИНЯТИЕ ИНВЕСТИЦИОННЫХ РЕШЕНИЙ НА ОСНОВЕ ТЕОРИИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ДИСКРЕТНЫМИ СИСТЕМАМИ

О.В. Павлов

На основе теории оптимального управления дискретными системами предложен общий подход к принятию инвестиционных решений. Сформулирована и решена, опираясь на принцип максимума Л.С. Понтрягина, задача оптимального управления инвестициями в дискретной постановке. Предложен численный алгоритм определения оптимального управления инвестициями. Получено аналитическое решение для случая постоянной рентабельности основных средств. Сформулированы критерии для принятия решения об инвестировании средств в реальные активы.

**Ключевые слова:** инвестиционное решение, оптимальное управление, инвестиции, принцип максимума Понтрягина, численный алгоритм, критерии.

## ВВЕДЕНИЕ

Под инвестиционным решением понимается выбор объектов инвестирования (реальных активов), времени начала и окончания инвестирования и объема необходимых капиталовложений. Для оценки инвестиционных решений применяются следующие критерии [1–5]: чистый приведенный (дисконтированный) доход  $NPV$  (*Net Present Value*), внутренняя норма доходности  $IRR$  (*Interval Rate of Return*), дисконтированный срок окупаемости  $DPP$  (*Discounted Payback Period*), индекс рентабельности  $PI$  (*Profitability Index*). Общепринято считать наилучшим критерий чистого приведенного дохода  $NPV$ . В настоящей статье для выбора инвестиционных решений применяется теория оптимального управления дискретными процессами [6–8]. Данный подход используется для первого качественного анализа целесообразности инвестирования в проект. После принятия инвестиционного решения на основе этого подхода необходимо более подробное обоснование экономической эффективности проекта.

Задачи управления инвестициями в непрерывной постановке рассматривались в работах [9–12]. В качестве критерия использовалась сумма дисконтированной прибыли предприятия за интервал планирования. Однако для практических экономических расчетов более предпочтительно решение задачи в дискретном виде.

## 1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДЕНЕЖНЫХ ПОТОКОВ ИНВЕСТИЦИОННОГО ПРОЕКТА

Промышленное предприятие является единственным участником инвестиционного проекта по производству нового вида продукции. Инвестиции осуществляются за счет собственных средств предприятия. Экономическая эффективность проекта оценивается в целом, и схема финансирования не учитывается. Рассматриваются денежные потоки от операционной (производственной) и инвестиционной деятельности [5]. В результате осуществления проекта предприятие работает на максимальную мощность и вся производимая продукция полностью продается. Считается, что денежный поток, генерируемый инвестиционным проектом в течение одного временного периода, имеет место в конце периода, т. е. имеет место постнумерандо. Предполагается, что рассматриваемая математическая модель денежных потоков отражает осредненную по отрасли ситуацию.

В качестве критерия оценки экономической эффективности инвестиционного проекта рассматривается чистый приведенный доход:

$$NPV = \sum_{t=0}^n \frac{FCF_t}{(1+r)^t}, \quad (1)$$

где  $t$  — номер периода инвестиционного проекта,  $FCF_t$  — свободный денежный поток инвестицион-

ного проекта (*Free Cash Flow*) в периоде  $t$ ,  $n$  — горизонт планирования, измеряемый числом периодов,  $r$  — ставка дисконтирования.

Горизонт планирования определяется как объективным фактором — жизненным циклом инвестиционного проекта, так и субъективным — дальновидностью менеджера, принимающего решение.

Свободный денежный поток инвестиционного проекта  $FCF_t$  в конце периода  $t$  определяется как разница денежных потоков от операционной  $OCF_t$  (*Operating Cash Flow*) и инвестиционной деятельности  $ICF_t$  (*Investment Cash Flow*) [1—5]:

$$FCF_t = OCF_t - ICF_t, \quad t = 0, n. \quad (2)$$

Денежный поток от операционной деятельности рассчитывается (здесь и далее предполагается  $t = 0, n$ ) как

$$OCF_t = REV_t - NOC_t - PT_t, \quad (3)$$

где  $REV_t$  — выручка (*Revenue*) от реализации произведенной продукции,  $NOC_t$  — чистые операционные издержки (*Net Operating Costs*),  $PT_t$  — налог на прибыль (*Profit Tax*).

Выручка

$$REV_t = P_t Q_t, \quad (4)$$

где  $P_t$  — цена продукции,  $Q_t$  — прогнозируемый объем продаж продукции.

Чистые операционные издержки включают в себя материальные затраты  $MC_t$  (*Material Costs*), заработную плату  $WS_t$  (*Wages and Salary*) и начисления на заработную плату  $WC_t$  (*Wages Charges*):

$$NOC_t = MC_t + WS_t + WC_t. \quad (5)$$

Материальные затраты

$$MC_t = Cm_t Q_t, \quad (6)$$

$Cm_t$  — материальные затраты на единицу продукции.

Фонд заработной платы

$$WS_t = w_t L_t, \quad (7)$$

где  $w_t$  — средняя ставка заработной платы персонала,  $L_t$  — численность производственного персонала.

Численность персонала рассчитывается по формуле

$$L_t = Q_t / l_t, \quad (8)$$

где  $l_t$  — норматив выпуска продукции средним работником за период  $t$ .

Начисления на заработную плату

$$WC_t = \tau_w WS_t, \quad (9)$$

где  $\tau_w$  — ставка единого социального налога.

Из формул (5)—(9) получим окончательное выражение для чистых операционных издержек

$$NOC_t = Q_t C_t, \quad (10)$$

где себестоимость  $C_t$  продукции (затраты на единицу продукции) рассчитывается по формуле

$$C_t = Cm_t + (1 + \tau_w) w_t / l_t. \quad (11)$$

Налог на прибыль вычисляется как

$$PT_t = \tau_c (REV_t - NOC_t - DEP_t), \quad (12)$$

где  $\tau_c$  — ставка налога на прибыль,  $DEP_t$  — амортизационные начисления (*Depreciation*).

Для расчета износа основных средств (внеоборотных активов) предприятия  $FA_t$  (*Fixed Assets*) применяется метод равномерного начисления амортизации:

$$DEP_t = \mu FA_t, \quad (13)$$

где  $\mu$  — норма амортизации,  $FA_t$  — стоимость основных средств предприятия в начале периода  $t$  до новых капиталовложений.

Все основные средства проекта полностью используются для производства продукции. В этом случае процесс производственной деятельности предприятия описывается производственной функцией Леонтьева:

$$P_t Q_t = f FA_t, \quad (14)$$

где  $P_t Q_t$  — стоимость прогнозируемого объема продаж продукции (выручка),  $f$  — фондоотдача основных средств, характеризующая применяемый технологический процесс для производства продукции.

Из формул (3), (4), (10), (12) и (13) получим

$$OCF_t = [(1 - \tau_c) f (P_t - C_t) / P_t + \tau_c \mu] FA_t.$$

Введем понятие рентабельности (доходности) основных средств (внеоборотных активов) проекта в форме денежного потока  $ROFA_t$  (*cash flow Return On Fixed Assets*):

$$ROFA_t = OCF_t / FA_t = (1 - \tau_c) f (P_t - C_t) / P_t + \tau_c \mu. \quad (15)$$

Экономический смысл рентабельности основных средств состоит в размере операционного денежного потока, приходящегося на единицу стоимости основных средств, используемых в проекте. Рентабельность основных средств зависит от цены и себестоимости продукции, фондоотдачи, ставки налога на прибыль и налогового щита, связанного с амортизацией  $\tau_c \mu$ . Рентабельность основных средств по денежному потоку по своему экономическому смыслу близка к понятию рентабельности инвестиций по денежным потокам [2, 4].

С учетом выражения (15) операционный денежный поток

$$OCF_t = ROFA_t FA_t. \quad (16)$$



Инвестиционный денежный поток  $ICF_t$  расходуется на капиталовложения в основные средства  $INV_t$  и изменение оборотного капитала  $\Delta NWC_t$  [1–5]:

$$ICF_t = \alpha_t INV_t + \Delta NWC_t, \quad (17)$$

где  $\alpha_t$  — процент от финансового ресурса предприятия  $INV_t$  в период  $t$ , необходимый для капиталовложений в проект. Если весь финансовый ресурс  $INV_t$  инвестируется в период  $t$  в рассматриваемый проект, то  $\alpha_t = 1$ , если не инвестируется, то  $\alpha_t = 0$ .

Согласно рекомендациям [5] чистый оборотный капитал (*Net Working Capital*) может быть рассчитан как процент от выручки:  $NWC_t = \gamma_t REV_t$ . Процент  $\gamma_t$  характеризует эффективность управления оборотным капиталом в периоде  $t$ .

Изменение потребности в финансировании оборотного капитала возникает при увеличении объема производства:  $\Delta NWC_t = \gamma_t (REV_{t+1} - REV_t)$ . С учетом формул (4) и (14) увеличение оборотного капитала

$$\Delta NWC_t = \gamma_t f (FA_{t+1} - FA_t) = \varphi_t \Delta FA_t. \quad (18)$$

Экономический смысл коэффициента  $\varphi_t$  состоит в количестве денежных средств, необходимым для инвестирования в прирост оборотного капитала при увеличении стоимости основных средств на одну денежную единицу для увеличения производства продукции.

Стоимость основных средств проекта в период  $t + 1$  определяется как сумма стоимости основных средств в предшествующий период  $t$  и капиталовложений за вычетом уменьшения стоимости основных средств в связи с их износом [1, 2]:

$$FA_{t+1} = FA_t - k_t FA_t + \alpha_t INV_t, \quad t = 0, n, \quad (19)$$

где  $k_t$  — коэффициент выбытия основных средств в периоде  $t$ .

В начальный период для основных средств задано начальное условие

$$FA_0 = FA_0. \quad (20)$$

Изменение стоимости основных средств определится из формулы (19):

$$\Delta FA_t = \alpha_t INV_t - k_t FA_t. \quad (21)$$

Подставим выражение (21) в формулу (18):

$$\Delta NWC_t = \varphi_t (\alpha_t INV_t - k_t FA_t). \quad (22)$$

С учетом формулы (22) инвестиционный денежный поток (17)

$$ICF_t = \alpha_t INV_t (1 + \varphi_t) - \varphi_t k_t FA_t. \quad (23)$$

Финансовый ресурс, необходимый для капиталовложений в проект, складывается из двух частей:

$$INV_t = INV_{1t} + INV_{2t}. \quad (24)$$

Первая часть необходима для увеличения основных средств из-за увеличения объема производства продукции, вторая — для восполнения основных средств из-за их износа. Они рассчитываются следующим образом:

$$INV_{1t} = \begin{cases} [Q_{t+1} - \max_{j=0,t} \{Q_j\}] P_{t+1} / f, \\ \text{если } Q_{t+1} > \max_{j=0,t} \{Q_j\}, \\ 0, \text{ если } Q_{t+1} \leq \max_{j=0,t} \{Q_j\}, t = 0, n, \end{cases} \quad (25)$$

$$INV_{2t} = \begin{cases} FA_t k_t, \text{ если } FA_t \geq FA_{t-1}, \\ \max_{j=0,t-1} \{FA_j\} - FA_t + FA_t k_t, \\ \text{если } FA_t < FA_{t-1}, t = 1, n. \end{cases} \quad (26)$$

Для начального периода  $INV_{20} = FA_0 k_0$ .

Выражение для чистого приведенного дохода (1) инвестиционного проекта с учетом формул (2), (6) и (23) примет вид:

$$NPV = \sum_{t=0}^n \frac{(ROFA_t + \varphi_t k_t) FA_t - \alpha_t INV_t (1 + \varphi_t)}{(1+r)^t}. \quad (27)$$

## 2. ПОСТАНОВКА И РЕШЕНИЕ ДИСКРЕТНОЙ ЗАДАЧИ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ ИНВЕСТИЦИЯМИ

Сформулируем проблему принятия инвестиционных решений как задачу оптимального управления дискретным процессом. В качестве управления примем процент  $\alpha_t$  от финансового ресурса предприятия  $INV_t$ , необходимого для капиталовложений в проект и рассчитываемого по формулам (24)–(26). Исходя из экономического смысла, на управление  $\alpha_t$  наложено ограничение

$$0 \leq \alpha_t \leq 1. \quad (28)$$

Сформулируем задачу оптимального управления: зная начальное состояние основных средств (20) проекта, необходимо выбрать такое допустимое управление (28) инвестициями для дискретного процесса (19), чтобы чистый приведенный доход (27) проекта принял максимальное значение.

Применим для решения задачи дискретный принцип максимума Л.С. Понтрягина [6–8]. Запишем гамильтониан

$$H_t = \left\{ \Psi_{t+1} - \frac{1 + \varphi_t}{(1+r)^t} \right\} \alpha_t INV_t + \Psi_{t+1} FA_t (1 - k_t) + \frac{(ROFA_t + \varphi_t k_t) FA_t}{(1+r)^t}, \quad t = 0, n.$$

Он линейно зависит от управления. В соответствии с принципом максимума [6–8] в каждой точ-

ке оптимальной траектории функция  $H_t$  достигает максимума относительно управления.

Из условия максимума гамильтониана найдем структуру оптимального управления:

$$\alpha_t^{\text{opt}} = \begin{cases} 1, & \text{если } \Psi_{t+1} - \frac{1 + \varphi_t}{(1+r)^t} \geq 0, \\ 0, & \text{если } \Psi_{t+1} - \frac{1 + \varphi_t}{(1+r)^t} < 0, \quad t = 0, n. \end{cases} \quad (29)$$

В работе [8] доказано, что для дискретных систем, линейных по фазовым координатам, принцип максимума справедлив — как необходимое и достаточное условие оптимальности. Таким образом, полученное условие (29) является условием оптимальности.

Для сопряженного уравнения [6, 7]

$$\Psi_t = \frac{\partial H_t}{\partial FA} = (1 - k_t)\Psi_{t+1} + \frac{ROFA_t + \varphi_t k_t}{(1+r)^t}, \quad t = n, 0, \quad (30)$$

на правом конце должно выполняться условие трансверсальности

$$\Psi_{n+1} = 0. \quad (31)$$

Из условий (29) и (31) следует: инвестирование в последний период  $n$  не производится,  $\alpha_n = 0$ . Теоретически, в связи со скачкообразным изменением рентабельности основных средств, связанным, например, с колебаниями цен, возможно изменение знака функции

$$\Psi_{t+1} - \frac{1 + \varphi_t}{(1+r)^t},$$

а следовательно, и управление может переключаться несколько раз. В частном случае, если функция отрицательна на всем интервале, оптимальным решением будет отказ от инвестирования в проект.

Анализируя исходное (19) и сопряженное (30) уравнения, структуру оптимального управления (29), можно сделать вывод: уравнения решаются независимо друг от друга. Если рассчитать сопряженные переменные от конечного периода до начального по уравнению (30), используя условие трансверсальности (31), то возможно определить оптимальное управление по формуле (29). А затем вычислить исходную переменную — стоимость основных средств от начального периода до конечного по уравнению (19), используя начальное условие (20).

Из анализа следует и другой вывод: оптимальное инвестиционное решение определяется рентабельностью основных средств по денежному потоку  $ROFA_t$ , ставкой дисконтирования  $r$ , горизонтом планирования  $n$ , коэффициентом выбытия основных средств  $k_t$ , коэффициентом  $\varphi_t$ , учитывающим

потребность в инвестировании при увеличении оборотного капитала.

Для определения оптимального управления инвестициями разработан следующий численный алгоритм, идея которого предложена в работе [12].

**Шаг 1.** Подготовка исходных данных проекта: прогнозируемых цен  $P_t$ , объемов продаж продукции  $Q_t$ , материальных затрат на единицу продукции  $Cm_t$ , числа сотрудников  $L_t$  и средней зарплаты сотрудников  $w_t$  для периодов от 0 до  $n$ .

**Шаг 2.** Расчет по формуле (11) себестоимости продукции  $C_p$ ,  $t = 0, n$ .

**Шаг 3.** Вычисление по формуле (15) рентабельности основных средств  $ROFA_t$ ,  $t = 0, n$ .

**Шаг 4.** Определение по формулам (24)—(26) объема финансового ресурса  $INV_t$ , необходимого для инвестирования в проект,  $t = 0, n$ .

**Шаг 5.** Расчет по формуле (30) сопряженной переменной  $\Psi_{t+1}$ ,  $t = n, 0$ .

**Шаг 6.** Нахождение оптимального управления по формуле (29) для периодов от 0 до  $n$ .

**Шаг 7.** Определение оптимальных капиталовложений  $\alpha_t INV_t$ ,  $t = 0, n$ .

**Шаг 8.** Расчет по формуле (19) стоимости основных фондов  $FA_t$ ,  $t = 0, n$ .

**Шаг 9.** Определение по формуле (16) операционного денежного потока  $OCF_t$ ,  $t = 0, n$ .

**Шаг 10.** Вычисление по формуле (23) инвестиционного денежного потока  $ICF_t$ ,  $t = 0, n$ ;

**Шаг 11.** Вычисление чистого приведенного дохода  $NPV$  по формуле (27).

Предложенный алгоритм был реализован в электронной таблице Excel. Численное моделирование показало его работоспособность и эффективность.

Если в результате расчетов оптимальное управление инвестициями на всей траектории будет равно нулю, то рассматриваемый проект необходимо отклонить. Возможна ситуация, когда в начальных периодах оптимальное управление инвестициями будет равно нулю, а затем из-за изменения экономических параметров проекта, например, роста цены на продукцию или снижения себестоимости, будет равна единице. В этом случае в результате расчетов определится оптимальное время начала осуществления инвестиционного проекта.

### 3. АНАЛИТИЧЕСКОЕ РЕШЕНИЕ ДИСКРЕТНОЙ ЗАДАЧИ ДЛЯ ПОСТОЯННОЙ РЕНТАБЕЛЬНОСТИ ОСНОВНЫХ СРЕДСТВ ПО ДЕНЕЖНОМУ ПОТОКУ

В частном случае, когда рентабельность основных средств инвестиционного проекта, коэффициент выбытия основных средств и коэффициент  $\varphi_t$  постоянны:  $ROFA_t = ROFA = \text{const}$ ,  $k_t = k = \text{const}$ ,



$\varphi_t = \varphi = \text{const}$ , возможно аналитическое решение дискретной задачи. Уравнение для сопряженной переменной (30) запишется в виде

$$\Psi_t = (1 - k)\Psi_{t+1} + \frac{ROFA + \varphi k}{(1 + r)^t}, \quad t = n, 0. \quad (32)$$

Условие трансверсальности  $\Psi_{t+1} = 0$ .

Так как формула (32) является рекуррентной, можно вывести выражение для сопряженной переменной через параметры проекта. Запишем сопряженную переменную в периоды  $t = n, n - 1, n - 2, n - 3, \dots$  по уравнению (32), используя условие трансверсальности (31):

$$\Psi_{n+1} = 0,$$

$$\Psi_n = \frac{ROFA + \varphi k}{(1 + r)^n},$$

$$\Psi_{n-1} = \frac{ROFA + \varphi k}{(1 + r)^{n-1}} \left[ 1 + \frac{1-k}{1+r} \right],$$

$$\Psi_{n-2} = \frac{ROFA + \varphi k}{(1 + r)^{n-2}} \left[ 1 + \frac{1-k}{1+r} + \left( \frac{1-k}{1+r} \right)^2 \right],$$

$$\Psi_{n-3} = \frac{ROFA + \varphi k}{(1 + r)^{n-3}} \left[ 1 + \frac{1-k}{1+r} + \left( \frac{1-k}{1+r} \right)^2 + \left( \frac{1-k}{1+r} \right)^3 \right].$$

Обобщая, запишем формулу для переменной  $\Psi_{t+1}$ :

$$\Psi_{t+1} = \frac{ROFA + \varphi k}{(1 + r)^{t+1}} \times \left[ 1 + \frac{1-k}{1+r} + \left( \frac{1-k}{1+r} \right)^2 + \left( \frac{1-k}{1+r} \right)^3 + \dots + \left( \frac{1-k}{1+r} \right)^{n-t-1} \right].$$

Выражение в скобках представляет собой сумму геометрической прогрессии со знаменателем прогрессии  $\frac{1-k}{1+r}$  и числом членов прогрессии  $n - t$ .

Применяя формулу для суммы геометрической прогрессии, получим следующее выражение для сопряженной переменной:

$$\Psi_{t+1} = \frac{ROFA + \varphi k}{r + k} \left[ \frac{(1 + r)^{n-t} - (1 - k)^{n-t}}{(1 + r)^{n-t}} \right]. \quad (33)$$

С учетом формулы (33) оптимальное управление (29) инвестициями примет вид:

$$\alpha_t^{\text{opt}} =$$

$$= \begin{cases} 1, & \text{если } \frac{ROFA + \varphi k}{r + k} \left[ \frac{(1 + r)^{n-t} - (1 - k)^{n-t}}{(1 + r)^{n-t}} \right] \geq 1 + \varphi, \\ 0, & \text{если } \frac{ROFA + \varphi k}{r + k} \left[ \frac{(1 + r)^{n-t} - (1 - k)^{n-t}}{(1 + r)^{n-t}} \right] < 1 + \varphi. \end{cases} \quad (34)$$

Функция

$$\frac{ROFA + \varphi k}{r + k} \left[ 1 - \left( \frac{1 - k}{1 + r} \right)^{n-t} \right]$$

монотонно убывающая, следовательно, возможно переключение управления не более одного раза в периоды от 0 до  $n$ :

$$\alpha_t^{\text{opt}} = \begin{cases} 1, & \text{если } 0 \leq t \leq t^{\text{KP}}, \\ 0, & \text{если } t^{\text{KP}} < t \leq n, \quad t = 0, n, \end{cases} \quad (35)$$

где  $t^{\text{KP}}$  — критический период проекта, период прекращения инвестиций.

Таким образом, оптимальная стратегия состоит в инвестировании финансовых средств в периоды от 0 до  $t^{\text{KP}}$  и в полном отказе от инвестирования после периода  $t^{\text{KP}}$ .

Из выражения (35) следует, что инвестирование в периоде  $t$  выгодно, если для рентабельности основных средств по денежным потокам выполняется условие:

$$ROFA \geq \frac{(1 + \varphi)(r + k)}{1 - \left( \frac{1 + r}{1 - k} \right)^{t-n}} - \varphi k.$$

Правая часть неравенства представляет собой критическое значение  $ROFA_t^{\text{KP}}$  рентабельности основных средств по денежным потокам, при котором инвестирование не выгодно, и условие (34) переписывается в виде:

$$\alpha_t^{\text{opt}} = \begin{cases} 1, & \text{если } ROFA > ROFA_t^{\text{KP}}, \\ 0, & \text{если } ROFA < ROFA_t^{\text{KP}}. \end{cases} \quad (36)$$

Полученное условие (36) может служить критерием для принятия решения об инвестировании средств в тот или иной проект. Если при заданном горизонте планирования  $n$  и ставки дисконтирования  $r$  рентабельность основных средств  $ROFA$  меньше критического значения, то инвестиционный проект следует отклонить.

Для предварительных расчетов, когда можно не учитывать выбытие основных средств ( $k = 0$ ) и инвестирование в увеличение оборотного капитала ( $\varphi = 0$ ), критическое значение рентабельности основных средств по денежным потокам принимает вид:

$$ROFA^{\text{KP}}(t, n, r) = \frac{r}{1 - (1 + r)^{t-n}}, \quad (37)$$

т. е. является функцией периода  $t$ , в котором производятся инвестиции, горизонта планирования  $n$  и ставки дисконтирования  $r$ . Из формулы (37) сле-

дует, что при бесконечном инвестиционном проекте  $n \rightarrow \infty$  критическое значение рентабельности основных средств по денежным потокам стремится к ставке дисконтирования  $r$ . Таким образом, для бесконечного инвестиционного проекта рентабельность основных средств должна быть не меньше ставки дисконтирования.

На рис. 1 представлены результаты расчетов критического значения рентабельности основных средств в зависимости от периода  $t$ , в котором производится инвестирование. Расчеты проводились для горизонта планирования  $n = 10$  и трех ставок дисконтирования  $r_1 = 10$ ,  $r_2 = 15$  и  $r_3 = 20$  %. Видно, что с увеличением периода  $t$  критическое значение рентабельности увеличивается до бесконечности в период  $t = n$ , а следовательно, обязательно наступит момент, когда инвестиции станут невыгодными.

Для начального периода  $t = 0$

$$\begin{aligned} ROFA^{KP}(0, n, r) &= \frac{r}{1 - (1+r)^{-n}} = \\ &= \left( \frac{1}{r} - \frac{1}{r(1+r)^n} \right)^{-1} = \frac{1}{B(n, r)}, \end{aligned}$$

где  $B(n, r)$  — коэффициент аннуитета,  $1/B(n, r)$  — коэффициент возврата капитала [2] (*capital recovery factor*).

Таким образом, для осуществления инвестиций в начальный период времени рентабельность основных средств должна быть не меньше коэффициента возврата капитала.

На рис. 2 приведены результаты расчетов критического значения рентабельности основных средств в начальный период  $t = 0$  в зависимости от

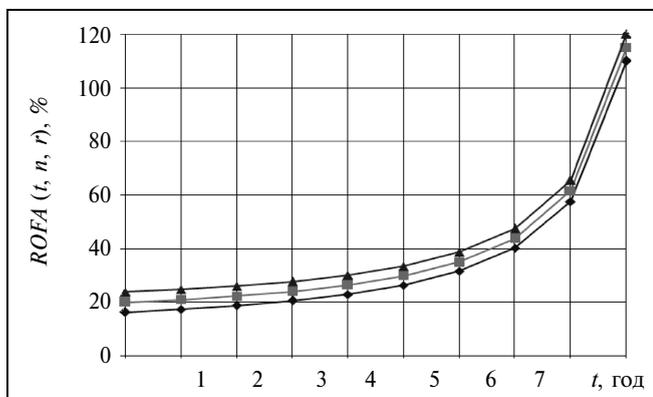


Рис. 1. Зависимость критического значения рентабельности основных средств от периода инвестирования:

◆ —  $r_1 = 10$  %; ■ —  $r_2 = 15$  %; ▲ —  $r_3 = 20$  %

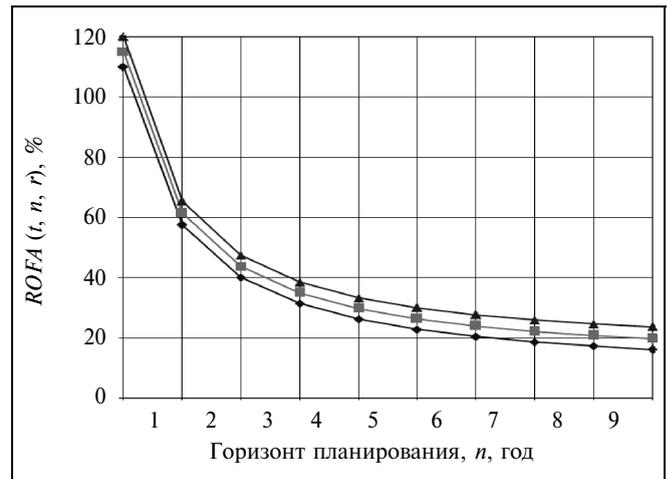


Рис. 2. Зависимость критического значения рентабельности основных средств от горизонта планирования:

◆ —  $r_1 = 10$  %; ■ —  $r_2 = 15$  %; ▲ —  $r_3 = 20$  %

горизонта планирования  $n$ . Расчеты проводились для трех ставок дисконтирования  $r_1 = 10$ ,  $r_2 = 15$  и  $r_3 = 20$  %. Видно, что с увеличением горизонта планирования  $n$  критическое значение рентабельности уменьшается, следовательно, чем больше горизонт планирования, тем меньшим может быть критическое значение рентабельности основных средств в начальный период.

Из равенства

$$\frac{RPFA + \varphi k}{r + k} \left[ 1 - \left( \frac{1-k}{1+r} \right)^{n-t} \right] - (1 + \varphi) = 0$$

в условии (34) определим период прекращения инвестиций:

$$t^{KP} = n + \log_{\frac{1+r}{1-k}} \left[ 1 - \frac{(1+\varphi)(r+k)}{ROFA + \varphi k} \right]. \quad (38)$$

Из экономического смысла следует, что выражение в скобках меньше единицы, поэтому логарифм отрицателен, и время прекращения инвестиций находится в интервале от 0 до  $n$ . Анализируя формулу (38), можно сделать следующий вывод: время прекращения инвестиций зависит от горизонта планирования: чем больший период времени предполагается осуществлять инвестиционный проект, тем дольше необходимо инвестировать в основные средства. Формула (38) математически определяет зависимость инвестиционного решения от дальновидности менеджера (жизненного цикла проекта). Период прекращения инвестиций также зависит от соотношения рентабельности основных средств  $ROFA$  и ставки дисконтирования  $r$ .



Для предварительных расчетов, когда можно не учитывать выбытие основных средств ( $k = 0$ ) и финансирование увеличения оборотного капитала ( $\varphi = 0$ ), период прекращения инвестиций

$$t^{KP} = n + \log_{1+r} \left[ 1 - \frac{r}{ROFA} \right].$$

### ЗАКЛЮЧЕНИЕ

Предложен общий подход к принятию инвестиционных решений, основанный на теории оптимального управления дискретными процессами. Сформулирована и решена задача оптимального управления инвестициями в дискретной постановке. Найдена структура оптимального управления инвестициями в общем виде. На основе полученного решения предложен численный алгоритм определения оптимального управления инвестициями, который реализован в электронной таблице Excel. Практические расчеты показали эффективность предложенного алгоритма для принятия инвестиционных решений.

Введено понятие рентабельности основных средств в виде денежного потока  $ROFA_r$ . Сделан вывод, что оптимальное инвестиционное решение определяется рентабельностью основных средств  $ROFA_r$ , ставкой дисконтирования, горизонтом планирования  $n$ , коэффициентом  $\varphi_r$ , учитывающего потребность в финансировании прироста оборотного капитала и коэффициентом выбытия основных средств.

Для случая постоянной рентабельности основных средств по денежным потокам получено аналитическое решение. Введено понятие критического значения рентабельности основных средств по денежным потокам, которое является функцией периода, в котором производятся инвестиции, горизонта планирования и ставки дисконтирования.

Найдены условия инвестирования в проект в период  $t$ : рентабельность основных средств по денежным потокам в период  $t$  должна быть больше критического значения. Это условие может использоваться как критерий для принятия решения об инвестировании в рассматриваемый проект. Из полученного условия найдена аналитическая формула для определения периода прекращения инвестиций. Для инвестиционного проекта с бесконечным жизненным циклом критическое значение рентабельности основных средств по денежным потокам должно быть не меньше ставки дисконтирования.

Дальнейшее развитие работы возможно в следующих направлениях:

- выбор наилучшего инвестиционного проекта из нескольких взаимоисключающих альтернатив;
- постановка и решение задачи оптимального управления финансированием инвестиционного проекта;
- исследование задачи оптимального управления инвестициями при изменяющейся конъюнктуре рынка: ценах и объемах продаж продукции, при различных стратегиях поведения конкурентов с применением теории игр;
- постановка и решение задачи оптимального управления инвестициями с учетом риска.

### ЛИТЕРАТУРА

1. Брейли Р., Майерс С. Принципы корпоративных финансов. — М.: ЗАО «Олимп — Бизнес», 1997. — 1120 с.
2. Бирман Г., Шмидт С. Капиталовложения: экономический анализ инвестиционных проектов. — М.: ЮНИТИ-ДАНА, 2003. — 631 с.
3. Бригем Ю., Гасперски Л. Финансовый менеджмент: Полный курс: В 2-х т. — СПб.: Экономическая школа, 1998.
4. Хелферт Э. Техника финансового анализа. — СПб.: Питер, 2003. — 640 с.
5. Виленский П.Л., Лившиц В.Н., Смоляк С.А. Оценка эффективности инвестиционных проектов: теория и практика. — М.: Дело, 2004. — 888 с.
6. Болтянский В.Г. Оптимальное управление дискретными системами. — М.: Наука, 1973. — 446 с.
7. Лагоша Б.А. Оптимальное управление в экономике. — М.: Финансы и статистика, 2003. — 192 с.
8. Розоноер Л.И. Принцип максимума Л.С. Понтрягина в теории оптимальных систем, III // Автоматика и телемеханика. — 1959. — № 12.
9. Соколовский Л.Е. Модели оптимального функционирования предприятия. — М.: Наука, 1980. — 175 с.
10. Горелик В.А., Кононенко А.Ф. Теоретико-игровые модели принятия решений в эколого-экономических системах. — М.: Радио и связь, 1982. — 144 с.
11. Косачев Ю.В. Экономико-математические модели эффективности финансово-промышленных структур. — М.: Логос, 2004. — 248 с.
12. Павлов О.В. Оптимальные стратегии привлечения финансовых ресурсов для развития фирмы в длительном периоде // Вестник Самарского государственного аэрокосмического университета. — 2004. № 1(5). — С. 147—152.
13. Павлов О.В. Численный алгоритм оптимального управления инвестициями промышленного предприятия // Экономические науки. Научно-информационный журнал. — 2009. — № 4(53). — С. 324—327.

Статья представлена к публикации членом редколлегии Р.М. Нижегородцевым.

**Павлов Олег Валерьевич** — канд. техн. наук, директор, Институт энергетики и транспорта Самарского государственного аэрокосмического университета им. С.П. Королёва,  
☎ (846) 950-54-77, ✉ pavlov@ssau.ru.

# СТРУКТУРНО-КЛАССИФИКАЦИОННАЯ МЕТОДОЛОГИЯ ОЦЕНКИ ЭФФЕКТИВНОСТИ ФУНКЦИОНИРОВАНИЯ ЖИЛИЩНО-КОММУНАЛЬНОГО ХОЗЯЙСТВА КРУПНОГО ГОРОДА (НА ПРИМЕРЕ МОСКВЫ)<sup>1</sup>

Ю.А. Дорофеев, А.С. Мандель

Рассмотрено применение методов классификационного анализа данных, многовариантной экспертизы и экспертно-статистической обработки информации к оценке эффективности и поддержки принятия решений в задачах управления жилищно-коммунальным хозяйством крупного города.

**Ключевые слова:** классификационный анализ данных, многовариантная экспертиза, экспертно-статистические методы, жилищно-коммунальное хозяйство, система оценки эффективности органов исполнительной власти.

## ВВЕДЕНИЕ

Оценка эффективности работы государственных и муниципальных органов — одна из наиболее сложных проблем теории и практики организационного управления. В отличие от коммерческих организаций, у которых есть простые экономические критерии эффективности типа прибыли, результаты деятельности государственных и муниципальных органов характеризуются состоянием объектов, которыми они управляют. Для оценки состояния этих объектов тоже существуют свои критерии, но эти критерии (например, готовность жилищного фонда к зиме, санитарное состояние дворовых территорий и т. п.), как правило, не поддаются непосредственному измерению. Чтобы оценить состояние объекта по этим критериям, необходимо проанализировать большое число первичных (непосредственно измеряемых) показателей. Для этого приходится применять достаточно сложные методы математической статистики и экспертного оценивания.

В настоящее время в Москве, как и во всех других крупных городах России, отсутствует единая система оценки эффективности функционирования жилищно-коммунального хозяйства (ЖКХ)

на уровне районов, административных округов и города в целом.

Более того, отсутствуют научно обоснованные методики оценки эффективности органов исполнительной власти практически во всех субъектах РФ. В июне 2007 г. вышел основополагающий документ в этой области — Указ Президента Российской Федерации «Об оценке эффективности деятельности органов исполнительной власти субъектов Российской Федерации» (от 26 июня 2007 г. № 825). В пп. 2.а этого документа Комиссии при Президенте РФ по вопросам совершенствования государственного управления и правосудия поручено «разработать и утвердить методику оценки эффективности деятельности органов исполнительной власти субъектов Российской Федерации». Хотя эта методика предназначена для оценки эффективности деятельности органов исполнительной власти субъектов РФ, совершенно очевидно, что одно из важнейших направлений деятельности органов исполнительной власти состоит в поддержании ЖКХ в надлежащем состоянии, без качественной оценки эффективности работы которого невозможна интегральная оценка эффективности деятельности исполнительной власти субъекта РФ в целом. Косвенным подтверждением этого служит «Перечень дополнительных показателей для оценки эффективности деятельности органов исполнительной власти субъектов Российской Федерации», утвержденный 18 июля 2007 г. упомянутой

<sup>1</sup> Работа выполнена при частичной финансовой поддержке РФФИ (проекты 08-07-00349-а, 10-07-00210-а, 09-07-00195-а).



Комиссией. В этом перечне около 18 % показателей либо непосредственно, либо косвенно относятся к ЖКХ.

Очень многие хозяйственные руководители пытаются объяснить невозможность качественной оценки эффективности деятельности органов управления отсутствием необходимой информации, но в подавляющем числе случаев это не соответствует действительности. Так, например, в Москве такие организации, как Объединение административно-технических инспекций Москвы (ОАТИ) и Мосжилинспекция (МЖИ) осуществляют регулярные проверки по срокам проведения и качеству выполнения работ. Кроме того, информация о нарушениях сроков и качества выполнения работ, об авариях и отключениях в системах электро-, газо- и водоснабжения поступает также и от граждан города. Однако вся эта информация не концентрируется в одном месте и глубоко не анализируется в целях подготовки качественных управленческих решений и оценки эффективности их реализации. В лучшем случае она используется для предъявления санкций отдельным подрядным организациям.

Такое положение затрудняет принятие эффективных решений по управлению ЖКХ города. Между тем, в системе ЖКХ задействованы огромные человеческие и финансовые ресурсы, поэтому несовершенство системы управления, в том числе оценки эффективности работы ее основных звеньев, приводит к существенным потерям.

Для того чтобы на основании собираемой ОАТИ, МЖИ, префектурами, управами и городскими диспетчерскими службами первичной информации можно было судить об эффективности системы управления ЖКХ города, необходимо разработать методику структурирования этой информации, формирования критериев эффективности, оценивания районов и округов по этим критериям и представления полученных таким образом оценок в сжатой и наглядной форме руководству городского хозяйства для принятия управленческих решений.

Для обоснования необходимости как разработки самой методики, так и включения отдельных разделов и конкретных требований к разрабатываемой методике, было проведено предварительное обследование существующей системы управления ЖКХ Москвы. Основные результаты обследования, а также аналитический и экспертный анализ этих результатов легли в основу методологии разработки системы оценки эффективности функционирования ЖКХ крупного города.

В настоящей работе описано применение методов классификационного анализа данных [1], включая методы автоматической классификации [2] и

экстремальной группировки параметров [3], а также методы многовариантной экспертизы [4] и экспертно-статистической обработки информации [5], для разработки концепции, алгоритмов и процедур мониторинга и оценки эффективности системы управления ЖКХ крупного города (на примере Москвы).

## 1. КОНЦЕПЦИЯ СИСТЕМЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ

Основные компоненты концепции разрабатываемой системы:

— формирование системы показателей, необходимых для оценки реальной эффективности основных уровней системы управления ЖКХ города, которые можно надежно собирать с заданной периодичностью;

— разработка схемы и процедур мониторинга этих показателей, определение носителей информации, средств связи, аппаратно-программного комплекса хранения и выдачи информации потребителям;

— разработка критериев оценки эффективности основных уровней управления ЖКХ города с использованием сформированной системы показателей;

— создание модели и методики оценки эффективности системы управления ЖКХ города на базе полученных показателей и критериев.

Концепция формировалась с учетом результатов экспертизы руководителей основных подразделений системы управления ЖКХ, которая проводилась на базе методологии коллективной многовариантной экспертизы [4, 6].

Как показало предварительное обследование, проведенное в Москве, существующая система управления ЖКХ и благоустройством крупного города может быть описана стандартной моделью управления слабоформализованным динамическим объектом с обратной связью в условиях сильного влияния человеческого фактора. Рассмотрим структурную схему такой модели (рис. 1).

Под объектом в модели понимается совокупность организаций, предприятий, подразделений и служб городского хозяйства, производящих работы или оказывающие услуги, необходимые для поддержания жилищно-коммунального хозяйства города в состоянии, определяемом соответствующими нормативно-правовыми актами федерального и регионального и муниципального уровня.

В блок управления входят:

— Департамент ЖКХ и благоустройства и другие подразделения комплекса городского хозяйства, которые по Положениям об их деятельности

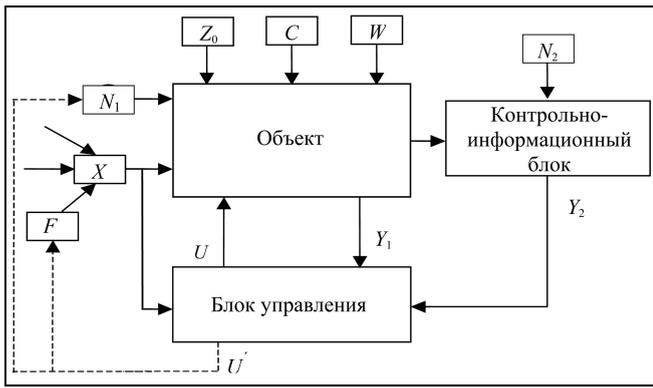


Рис. 1. Структурная схема модели управления ЖКХ и благоустройством на этапе обследования

наделены функциями принятия решений и выработки управляющих воздействий на соответствующие звенья управляемого объекта;

— другие подразделения Правительства Москвы, ответственные за отдельные направления деятельности ЖКХ и по благоустройству города (финансово-экономические, топливно-энергетические, экологические подразделения, наука, промышленное производство и др.);

— подразделения префектур административных округов и управ районов, ответственных за работу подразделений объекта, обслуживающих соответствующие территории.

Под контрольно-информационным блоком понимается совокупность организаций и подразделений Правительства Москвы, наделенных в соответствии с Положениями об их деятельности и другими нормативно-правовыми актами полномочиями и ответственностью за сбор достоверной информации о функционировании Объекта и своевременное представление ее в заданном формате в структуры блока управления.

На функционирование объекта влияют следующие факторы:

$X$  — неуправляемые или слабоуправляемые в краткосрочном режиме входные параметры, в том числе: плановое финансирование; планы по строительству, капитальному ремонту, реконструкции и текущему ремонту объектов ЖКХ и городской инфраструктуры; плановое материально-техническое обеспечение оборудованием, приборами, подвижным составом и средствами механизации, противогололедными реагентами и пр.; кадровое обеспечение и пр.;

$F$  — формирование адресных потоков финансирования для подразделений объекта — пример слабоуправляемых в краткосрочном режиме входных параметров;

$N_1$  — нормативно-правовая база функционирования объекта (ограничения);

$Z_0$  — начальные условия; важно учитывать, что они могут существенно отличаться для различных подразделений объекта, (так, например, число домов в аварийном или неудовлетворительном состоянии в Центральном административном округе Москвы почти в 12 раз больше, чем в следующем за ним по этому показателю Юго-Восточном административном округе);

$C$  — система целей, на достижение которых направлено функционирование объекта.

$W$  — неуправляемые и слабо прогнозируемые воздействия на объект (аналог случайного шума); в качестве примеров можно отметить погодно-климатические, природные, техногенные, экономические (кризис платежно-банковской системы, банкротства подрядных организаций, забастовки и пр.) воздействия;

$U$  — управляющие воздействия, поступающие на объект из блока управления (административные, финансовые, материально-технические, кадровые, в том числе связанные с материальным стимулированием или наказанием);

$U'$  — управляющие воздействия, поступающие из блока управления, которые носят среднесрочный или долгосрочный характер (стратегические); в качестве примера на рис. 1 показаны воздействия такого рода на нормативную базу и на планы финансирования различных подразделений объекта;

$Y_1$  — значения показателей, характеризующих текущее функционирование объекта; эта информация поступает в блок управления непосредственно с объекта;

$Y_2$  — значения показателей, характеризующих текущее функционирование объекта, но поступающие из контрольно-информационного блока;

$N_2$  — нормативно-правовая база функционирования контрольно-информационного блока.

**Основные выводы по результатам предварительного обследования.** В ходе предварительного обследования выяснилось, что наибольшие трудности в смысле управления представляет анализ первичных показателей  $Y_1$  и  $Y_2$  и их использование для оценки эффективности функционирования как основных служб и подразделений системы ЖКХ города, так и самой системы в целом.

Первый вопрос, который здесь возникает — на каком уровне системы ЖКХ города следует осуществлять этот анализ? Источниками первичной информации служат подразделения следующих организаций: префектуры административных округов; управы районов; МЖИ; ОАТИ; поставщики электрической и тепловой энергии, горячей и хо-



лодной воды и пр.; органы управления на уровне города, административного округа и района, работающие с обращениями и жалобами жителей города.

Один из важных результатов предварительного обследования, сформулированный после анализа результатов экспертных оценок, полученных в ходе проведения многовариантной экспертизы по этому вопросу, состоит в том, что в смысле баланса информативности (для оценки эффективности системы управления) и трудоемкости сбора информации *наиболее приемлемым является уровень района*.

Исторически в Москве сложилась ситуация, когда значительная часть первичной информации, необходимой для системы управления ЖКХ и собираемой на уровне района, формируется в окружных подразделениях ОАТИ и МЖИ.

**Структура и состав первичной информации (первичных показателей).** В ходе предварительного обследования было выяснено, что в состав первичной информации о состоянии жилищного фонда, территорий жилых кварталов и объектов внешнего благоустройства, находящихся на территории жилых кварталов, по каждому району Москвы должны входить сведения о показателях трех типов: технических, характеризующих социальную обстановку и экономических (рис. 2).

Структура системы первичных показателей представлена на рис. 2.

Технические показатели, собираемые рядом организаций и, прежде всего, Мосжилинспекцией и ОАТИ, прямо или косвенно характеризуют состояние системы ЖКХ в городе. В их число в настоящее время входят 49 показателей.

Выявлено 12 показателей, характеризующих социальную обстановку в связи с состоянием жилищного фонда, территорий жилых кварталов и объектов внешнего благоустройства. Кроме того, к числу косвенных показателей такого типа относятся сведения о травматизме населения Москвы (особенно зимой), которые поступают из Департамента здравоохранения. Эти показатели в определенной степени характеризуют состояние покрытия тротуаров, стоянок общественного транспорта, подходов к торговым и другим организациям, качество очистки этих объектов от снега и наледи (зимой), освещенность этих объектов в темное время суток и пр.

К числу экономических показателей относятся различные сведения о выделенных и освоенных объемах финансовых средств на поддержание и улучшение состояния системы ЖКХ, которые поступают из Департамента ЖКХ и благоустройства, Департамента экономики и финансов, Контроль-

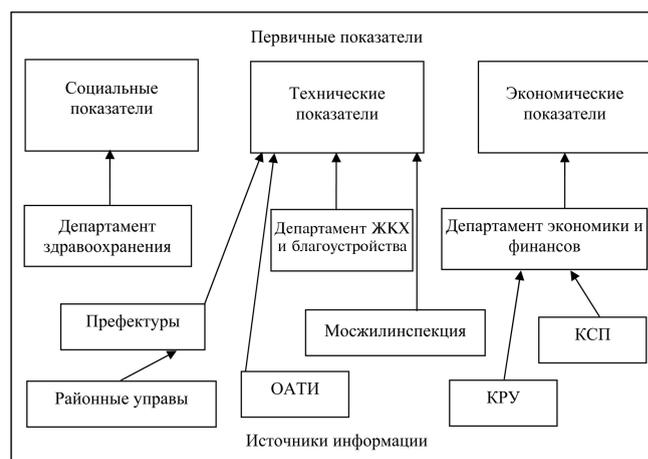


Рис. 2. Структура первичной информации

но-счетной палаты (КСП) и Контрольно-ревизионного управления (КРУ).

Большое значение придается работе с обращениями и жалобами, поступающими от жителей Москвы. Особое внимание уделяется повторным жалобам, а также обращениям по поводу аварийных и чрезвычайных ситуаций, которые ставятся на специальный контроль.

Первичная информация о ЖКХ города, содержащаяся в соответствующих формах ОАТИ и МЖИ, а также в данных, поступающих от диспетчерских служб города, удовлетворяет требованиям к надежности и периодичности. Она достаточно полно отражает реальное состояние объекта. Поэтому, по крайней мере, на первом этапе нет необходимости вводить новые показатели, тем более что введение новых показателей сопряжено с большими организационно-техническими трудностями (согласование в десятках организаций, получение разрешений и т. д.) и финансовыми затратами.

Подразделения ОАТИ, МЖИ и диспетчерские службы оснащены техническими средствами, позволяющими передавать собираемую первичную информацию в вышестоящие организации в электронном виде. Однако она передается в разные организации и не концентрируется в одном месте. Кроме того, информация накапливается в виде совокупности первичных форм учета и не усредняется по месяцам и кварталам.

Несмотря на полноту и достаточную надежность первичной информации, ее нельзя непосредственно использовать для оценки эффективности и поддержки принятия управленческих решений — даже в том случае, если будет обеспечена ее концентрация, усреднение и накопление в одном месте. Причина в ее большом объеме. Одни

только формы ОАТИ и МЖИ содержат данные по 60 показателям, а с учетом данных, поступающих в диспетчерские службы, число первичных показателей возрастает до 80 и более. Эти данные собираются по 124 районам города. Таким образом, первичная информация представляет собой таблицу (матрицу) размером не менее 124×80. Даже при использовании обычных компьютерных методов статистической обработки проанализировать такой массив данных и сделать какие-то выводы невозможно без специальной методики анализа больших массивов информации. Для того чтобы первичную информацию, представленную в виде таблиц большого размера, можно было использовать для целей управления, необходимо разработать методику структуризации такой информации, ее сжатия и представления в виде, удобном для лиц, принимающих решения (ЛПР).

**Необходимость структурного анализа первичной информации.** В разработанной концепции указанную структуризацию предлагается осуществить по двум направлениям — структуризация первичных показателей и структуризация оцениваемых объектов (т. е. районов города).

Структуризация первичных показателей необходима для того, чтобы на базе этих показателей сформировать небольшое число (5—10) достаточно информативных «интегральных» показателей, которые с минимальной потерей информации в более наглядной форме представляют информацию о состоянии объекта. Такие интегральные показатели можно будет использовать в качестве критериев оценки эффективности и качества проведения работ в системе ЖКХ. Как показывает практика, сформировать такие критерии чисто экспертным путем не удастся [2, 4]. Мнения экспертов, какие показатели более, а какие менее информативны, какие им следует приписать веса и т. п., часто расходятся. Здесь более надежные результаты дают специальные математические методы типа экстремальной группировки параметров или факторного анализа [1, 3, 4].

Поскольку в Москве более 120 районов, то даже при небольшом числе агрегированных показателей (критериев) информация остается трудно обозримой. Поэтому для формирования итоговых оценок необходима структуризация множества районов, т. е. разбиение этого множества на классы однотипных (в смысле сформированных критериев) районов. Тогда каждый район (соответственно административный округ) можно будет охарактеризовать не только количественно (набором значений критериев), но и качественно (принадлежностью к определенному классу). Это сделает оценку более наглядной и удобной для принятия

управленческих решений. При числе критериев больше двух такую структуризацию невозможно выполнить вручную. В разработанной концепции для этой цели предлагается применять специальные методы многомерной автоматической классификации [1, 2, 4].

На работу системы ЖКХ и благоустройства влияет множество не поддающихся формализации факторов, не все из которых можно учесть набором первичных показателей. Поэтому процедура формирования оценок должна быть человеко-машинной, т. е. допускать корректировку с учетом мнений экспертов. Для такой корректировки в разработанной концепции предусмотрено применять экспертно-классификационные [4, 6, 7] и экспертно-статистические [5, 8, 9] процедуры анализа данных.

Таким образом, в рамках концепции основное назначение разрабатываемой методики — структуризация первичной информации в целях представления ее в виде, пригодном для принятия управленческих решений. Это предполагает, прежде всего, структуризацию множества первичных показателей и формирование на их основе небольшого числа критериев; далее, структуризацию (классификацию) районов города и формирование на ее базе набора качественных оценок; наконец, формирование на этой основе качественных оценок эффективности работы соответствующих органов исполнительной власти.

---

## 2. ЭКСПЕРТНО-КЛАССИФИКАЦИОННЫЕ МЕТОДЫ СТРУКТУРИЗАЦИИ ПЕРВИЧНЫХ ПОКАЗАТЕЛЕЙ И ОБЪЕКТОВ ОЦЕНКИ

---

Как уже говорилось, в рамках разработанной концепции для структуризации первичных показателей с последующим формированием информативных параметров (критериев) эффективности, а также структуризации множества оцениваемых объектов (районов города) применяются апробированные методы классификационного анализа и экспертизы (экстремальная группировка, автоматическая классификация, многовариантная экспертиза, экспертно-статистическое оценивание и др.).

### 2.1. Методы структуризации первичных показателей

Воспользуемся для решения проблемы структуризации первичных показателей методами экстремальной группировки параметров [1—3]. Экстремальная группировка параметров предназначена для решения проблемы сокращения числа исходных (первичных) показателей функционирования сложной системы в целях выделения относительно



небольшого числа наиболее информативных, агрегированных показателей (факторов), которые предъявляются ЛПР для оценки эффективности управления соответствующей системой. Выделяемые в результате группировки агрегированные показатели обладают тем свойством, что они содержат в себе всю существенную информацию о характере происходящих в системе процессов и могут быть охарактеризованы функциональными связями с собираемыми в результате мониторинга первичными показателями, например, быть взвешенными суммами каких-либо первичных показателей.

В такой постановке первичные показатели рассматриваются как случайные величины, отдельные значения которых (реализации) становятся известными после их фиксации в системе мониторинга первичных показателей.

В дальнейшем коэффициент корреляции (или ковариации)  $\rho_{x, y}$  двух случайных величин  $x$  и  $y$  (первичных показателей) будем обозначать так:  $\rho_{x, y} = (x, y)$ , подчеркивая этим обозначением тот факт, что коэффициент корреляции может пониматься как скалярное произведение случайных величин  $x$  и  $y$ . Для дисперсии случайной величины  $x$  будем применять обозначение  $\rho_{x, x} = (x, x) = x^2$ . Множество первичных показателей (случайных величин)  $x^{(1)}, x^{(2)}, \dots, x^{(k)}$  разбито на непересекающиеся группы  $A_1, A_2, \dots, A_s$  и заданы случайные величины  $f_1, f_2, \dots, f_s$  такие, что  $f_j^2 = 1, j = 1 \div s$ , которые будем называть факторами. Введем в рассмотрение функционал

$$J^* = \sum_{x_i \in A_1} (x^{(i)}, f_1)^2 + \sum_{x_i \in A_2} (x^{(i)}, f_2)^2 + \dots + \sum_{x_i \in A_s} (x^{(i)}, f_s)^2. \quad (1)$$

Тогда задача экстремальной группировки показателей ставится как задача максимизации функционала (1) как по разбиению показателей на множества  $A_1, A_2, \dots, A_s$ , так и по выбору случайных величин  $f_1, f_2, \dots, f_s, f_j^2 = 1, j = 1 \div s$ .

Максимизация функционала (1) соответствует интуитивному требованию такого разбиения множества показателей, когда в одну группу попадают наиболее «близкие» между собой показатели. Действительно, при максимизации функционала (1) для каждого фиксированного набора случайных величин  $f_1, f_2, \dots, f_s$  в одну (например,  $l$ -ю группу) будут попадать такие показатели, которые наиболее «близки» к величине  $f_l$ ; в то же время

среди всех возможных наборов случайных величин  $f_1, f_2, \dots, f_s$  будет отбираться такой набор, что каждая из величин  $f_l$  в среднем наиболее «близка» ко всем показателям из своей группы.

В настоящей работе методология структуризации первичных показателей и отбора информативных параметров была реализована в составе комплексного алгоритма автоматической классификации [2], предназначенного для решения задач анализа и принятия решений в крупномасштабных системах управления.

## 2.2. Методы структуризации множества оцениваемых объектов

Для решения проблемы структуризации (классификации) объектов системы ЖКХ по набору выделенных информативных показателей (факторов) функционирования системы управления ЖКХ в работе применялись методы автоматической классификации.

Формальная постановка задачи автоматической классификации основана на введении в рассмотрение критерия качества классификации (функционала), зависящего от конкретного разбиения пространства  $X$  на области, экстремум которого соответствует интуитивному представлению о разбиении пространства  $X$  на «компактные» области. В работе для этой цели использовался функционал средней по классам близости точек в классах [2, 3]:

$$J_1 = \sum_{i=1}^r \frac{n_i}{n} K(A_i, A_i). \quad (2)$$

Здесь через  $K(A_i, A_i)$  обозначена средняя близость точек в классе  $A_i$

$$K(A_i, A_i) = \frac{2}{n_i(n_i - 1)} \sum_{i=1}^{n_i} \sum_{j>i} K(x_i, x_j),$$

где  $K(x_i, x_j)$ , — потенциальная функция, определяющая меру близости точек  $x_i$  и  $x_j$ ,  $n_i$  — число точек в классе  $A_i$ ,  $n$  — общее число точек. В работе при практических расчетах функция  $K(x_i, x_j)$ , использовалось в виде  $K(x_i, x_j) = 1/(1 + \alpha R^p(x_i, x_j))$ , где через  $R(x_i, x_j)$ , обозначено евклидово расстояние между точками  $x_i$  и  $x_j$  в пространстве параметров  $X$ ,  $\alpha$  и  $p$  — настраиваемые параметры алгоритма [2].

Требуется разбить пространство  $X$  на  $r$  областей (а при одномерной классификации — ось значений показателя на  $r$  интервалов) таким образом, чтобы доставить максимум функционалу (2). Для решения так поставленной задачи структуризации

объектов в работе используется комплексный алгоритм автоматической классификации [2].

В экспертно-классификационных методах исследования крупномасштабных систем управления используются не только результаты автоматической классификации в многомерном пространстве. В рамках методологии построения хорошо интерпретируемых классификаций [7] разработаны алгоритмы оптимальной одномерной классификации объектов на оси каждого из информативных показателей. Выделенные в результате такой классификации области на шкале значений каждого из этих показателей становятся оценками (в баллах) качества (эффективности управления) по данному показателю рассматриваемого объекта ЖКХ (например, района Москвы).

### 3. МЕТОДЫ ЭКСПЕРТНО-СТАТИСТИЧЕСКОЙ КОРРЕКТИРОВКИ РЕЗУЛЬТАТОВ

В результате решения задачи структурной обработки  $k$  первичных показателей  $x^{(1)}, x^{(2)}, \dots, x^{(k)}$  формируется список агрегированных (интегральных) информативных показателей — факторов  $f_1, f_2, \dots, f_s$ , число которых  $s$  существенно меньше числа  $k$  первичных показателей. На базе этих наборов первичных показателей и факторов формируется  $S$  критериев  $F_j$  эффективности функционирования системы ЖКХ города. Для этой цели используются экспертно-классификационные процедуры [4]. Значения критериев  $F_j$  оцениваются в первично выбранных шкалах (например, балльных). При этом указываются формулы (модели) связи между критериями и системой первичных показателей и факторов. В качестве таких моделей обычно используются простые линейные модели вида

$$F_j = \sum_{i=1}^K \beta_{ij} y_i, \quad j = 1, \dots, S,$$

где  $y_i = x^{(i)}$  для  $i = 1, \dots, k$  и  $y_i = f_i$  для  $i = k + 1, \dots, K$ , здесь  $K = k + s$ . Коэффициенты  $\beta_{ij}$  называются *силами связи*. Расстояния (меры близости)  $R_{lp}^{(j)}$ ,  $j = 1, 2, \dots, S$ ,  $l \neq p$ , между классами, выделенными при балльном ранжировании каждого из критериев — эти расстояния называются *степенями различимости балльных оценок по  $j$ -му критерию*. На первом этапе разделение значений критериев осуществляется по четырем рангам, т. е.  $l, p = 1, 2, 3, 4$ .

Для уточнения результатов агрегирования и сформированного в результате списка показателей (критериев) эффективности функционирования

системы ЖКХ города в работе используется методология экспертно-статистической обработки информации [5, 8—10].

На следующем этапе списки и балльные значения критериев поступают в экспертные комиссии для проведения многовариантной экспертизы в целях выработки вариантов решений по сравнительной оценке объектов ЖКХ (на уровне административных округов и районов). Затем эти варианты передаются ЛПР для окончательной оценки (ранжировки) объектов ЖКХ (на уровне административных округов). Организуется процедура экспертно-статистической обработки результатов структуризации, суть которой состоит в том, что эксперты получают право:

- изменять список критериев (введение новых критериев, удаление предложенных критериев);

- изменять модели связи критериев с первичными показателями и факторами (на уровне предложений о включении в модель или исключении из нее тех или иных первичных показателей и (или) факторов). При этом эксперты не могут предлагать изменить сами значения коэффициентов (*сил связи*);

- корректировать значения предложенных критериев (в баллах);

- изменять в сторону увеличения или уменьшения число выделенных по каждому из критериев классов (*диапазон шкалы* каждого из критериев).

Для этой цели организуется интерактивная процедура взаимодействия экспертов с системой поддержки принятия решений (СППР), в рамках которой любое из предложений экспертов по внесению указанных вариантов изменений система «комментирует», предлагая вниманию экспертов (на основе имеющейся в базе данных системы статистической информации) оценки последствий предлагаемых ими изменений [5]. Получив в свое распоряжение комментарии системы, эксперты имеют право дезавуировать «возражения» СППР и либо подтвердить некоторые (или все) из предложенных ими изменений, либо согласиться с ними, отказавшись от части (или всех) предлагаемых ими изменений. В случае подтверждения всех или части изменений СППР осуществляет пересчет решения с учетом всех подтвержденных изменений и формирует новые значения критериев в баллах, которые снова предлагаются вниманию экспертов. Процедура продолжается до тех пор, пока эксперты не откажутся от внесения каких бы то ни было изменений. При отказе экспертов от всех предложенных ими изменений сформированная в результате система критериев и сформированных СППР передается ЛПР для выработки окончательного решения.



Лицо, принимающее решение, имеет право на внесение всех упомянутых изменений, вплоть до требования агрегирования всех критериев в один единственный критерий оценки эффективности работы объектов ЖКХ (на уровне административных округов) в целях их строгого ранжирования.

### ЗАКЛЮЧЕНИЕ

Представлено краткое описание концепции системы оценки эффективности и поддержки принятия решений по управлению жилищно-коммунальным хозяйством крупного города (на примере Москвы). Для решения задач структуризации первичных показателей и объектов оценки (районов и административных округов города), а также корректировки сформированной в результате системы критериев эффективности предложена методология, основанная на экспертно-классификационных и экспертно-статистических алгоритмах обработки информации.

В настоящее время с использованием результатов структурно-классификационной экспертизы прорабатываются критерии, модели и научно-обоснованная методика оценки эффективности работы ЖКХ Москвы на уровне района, административного округа и города в целом.

### ЛИТЕРАТУРА

1. Бауман Е.В., Дорофеюк А.А. Классификационный анализ данных // Тр. междунар. конф. по проблемам управления / ИПУ РАН. — М. 1999. — Т. 1. — С. 62—67.
2. Дорофеюк Ю.А. Комплексный алгоритм автоматической классификации и его использование в задачах анализа и принятия решений // Таврический вестник информатики и математики. — 2008. — № 1. — С. 171—177.

3. Браверман Э.М., Мучник И.Б. Структурные методы обработки эмпирических данных. — М.: Наука, 1983. — 464 с.
4. Дорофеюк Ю.А., Гольдовская М.Д., Покровская И.В. Экспертно-классификационный анализ данных в задаче оценки эффективности функционирования крупномасштабных систем управления // Таврический вестник информатики и математики. — 2008. — № 2. — С. 159—165.
5. Мандель А.С. Экспертно-статистические методы обработки информации в интегрированных системах управления производством и технологическими процессами // Проблемы управления. — 2006. — № 6. — С. 55—59.
6. Дорофеюк А.А., Покровская И.В., Чернявский А.Л. Экспертные методы анализа и совершенствования систем управления // Автоматика и телемеханика. — 2004. — № 10. — С. 172—188.
7. Дорофеюк Ю.А., Дорофеюк А.А., Чернявский А.Л. Построение хорошо интерпретируемых классификаций — методология и алгоритмы // Управление развитием крупномасштабных систем (MLSD'2008): Тр. II Междунар. конф. / ИПУ РАН. — М., 2008. — С. 164—170.
8. Беляков А.Г., Мандель А.С. Прогнозирование временных рядов на основе метода аналогов (элементы теории экспертно-статистических систем). — М.: Институт проблем управления, 2002. — 60 с.
9. Беляков А.Г., Мандель А.С. Анализ достоверности выводов, формируемых с помощью экспертно-статистических систем. — М.: Институт проблем управления, 2002. — 64 с.
10. Классификационные алгоритмы оценки эффективности и поддержки принятия решений в задачах управления ЖКХ мегаполиса Москва / А.С. Мандель, А.А. Дорофеюк, А.Л. Чернявский, Д.В. Лифшиц // Таврический вестник информатики и математики. — 2008. — № 2. — С. 42—48.

Статья представлена к публикации членом редколлегии А.Д. Цвиркуном.

Дорофеюк Юлия Александровна — науч. сотрудник,  
☎ (495) 334-75-40, ✉ tigress86@bk.ru,

Мандель Александр Соломонович — д-р техн. наук,  
зав. лабораторией, ☎ (495) 334-89-69, ✉ manfoon@ipu.ru,

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

### НОВАЯ КНИГА

Чобану М. Многомерные многоскоростные системы обработки сигналов. — М.: Техносфера, 2009. — 480 с.  
ISBN 978-5-94836-233-5.

Рассмотрены многомерные (ММ) многоскоростные системы для обработки ММ цифровых сигналов. Впервые на русском языке систематически изложены теория и методы неразделимой обработки ММ сигналов, приведено все необходимое для разработки ММ многоскоростных систем, начиная с фундаментальных результатов из теории цифровой обработки ММ сигналов и заканчивая алгоритмами, программным и аппаратным обеспечением.

Важная особенность монографии — применение математических пакетов MATLAB, MAPLE, Singular и др., а также программ, написанных на языке С, на протяжении всего изложения. Приведены результаты реализации разработанных неразделимых операторов на основе процессоров общего назначения, сигнальных процессоров фирмы «Texas Instruments» и графических процессоров (GPU) фирмы «nVidia».

Приобрести книгу можно через Интернет <http://www.dom-knigi.ru/book.asp?Art=316984&CatalogID=7> или обратившись по e-mail: [tchobanou@gmail.com](mailto:tchobanou@gmail.com)

# МОДЕЛИРОВАНИЕ КОНФЛИКТНЫХ СИТУАЦИЙ С НЕСОГЛАСОВАННЫМИ ПРЕДСТАВЛЕНИЯМИ У АГЕНТОВ НА ОСНОВЕ ИГР НА ЛИНЕЙНЫХ КОГНИТИВНЫХ КАРТАХ

С.Г. Куливец

Рассмотрены две теоретико-игровых модели конфликтов между агентами в слабоструктурированной ситуации. В качестве представления знаний агентов о ситуации используются линейные когнитивные карты. Когнитивные карты у различных агентов могут отличаться.

**Ключевые слова:** теория игр, анализ ситуаций с использованием когнитивных карт, слабоструктурированные ситуации, равновесие Нэша.

## ВВЕДЕНИЕ

В некоторых задачах комплексного управления в сферах, связанных с жизнью общества (в социально-экономической, политической и др.), объект управления оказывается слабо структурированным. Понятие слабо структурированности подробно рассматривается в работе [1]. Одно из его свойств заключается в следующем: ситуация слабо структурированная, если ее основные параметры носят качественный, а не количественный характер, и их значения представляют собой субъективные оценки экспертов. Для решения задач управления слабоструктурированными ситуациями пользуются когнитивными картами. *Когнитивная карта* — это одна из возможных моделей представления знаний эксперта (или группы экспертов) о ситуации, описанная в виде взвешенного ориентированного графа. Вершины когнитивной карты соответствуют факторам, в терминах которых описывается ситуация. Фактор можно трактовать как переменную, например «Обороноспособность государства», которая может принимать различные значения, такие как «высокая», «низкая» и т. п. Взвешенная дуга трактуется как непосредственное причинно-следственное влияние одного фактора на другой.

Задачи анализа ситуаций на основе когнитивных карт можно разделить на два типа: *статические* и *динамические* [2]. Статический анализ — это анализ текущей ситуации, заключающийся в вы-

делении и сопоставлении непосредственных и опосредованных причинно-следственных путей влияния одних факторов на другие. Динамический анализ — это генерация и анализ возможных сценариев развития ситуации во времени. Далее мы будем рассматривать только задачу динамического анализа когнитивной карты. В работах [3, 4] подробно рассматриваются задачи, связанные с динамическим анализом когнитивных карт. Базовой моделью при рассмотрении этих задач служит модель импульсных процессов, предложенная в книге [5]. В работе [3] описываются решения обратной задачи для линейных моделей при задании фиксированной и нефиксированной целей управления с ограниченными ресурсами. Под фиксированной целью управления понимается задание численных значений, соответствующих лингвистическим переменным, определяющих требуемые состояния целевых факторов. Под *нефиксированной целью управления* понимается обеспечение желательных направлений изменения всех целевых факторов. Поиск решения задач управления для линейных моделей в общем случае осуществляется с помощью различных оптимизационных методов. По-видимому, первой работой, в которой было приведено общее описание теоретико-игровой модели взаимодействия нескольких агентов в динамической системе, представленной в виде когнитивной карты ситуации, была работа [6]. В ней были перечислены основания для классификации теоретико-игровых моделей конфликта на когнитивной карте и

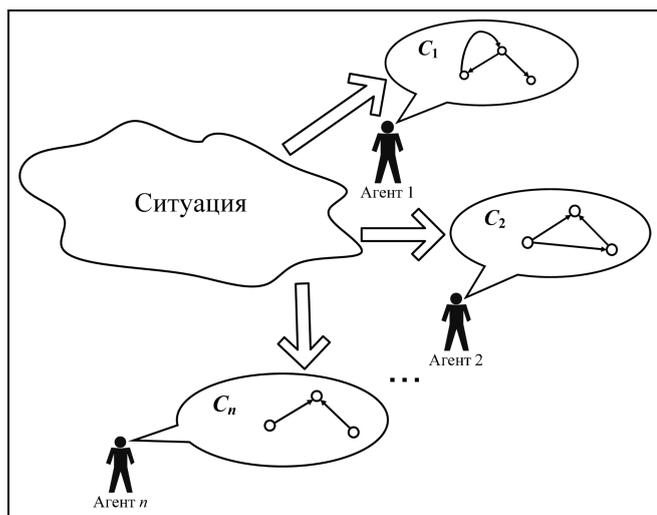


Рис. 1. Представления агентов о ситуации, формализованные в виде когнитивных карт

рассмотрена одна игра в нормальной форме на линейной когнитивной карте. Описание этой игры, в частности, послужило прямым прообразом для модели, рассмотренной в § 3 настоящей работы.

В настоящей статье рассматривается ситуация взаимодействия нескольких агентов, каждый из которых обладает в общем случае собственной когнитивной картой (рис. 1). В отличие от модели, рассмотренной в работе [6], не только цели агентов могут не совпадать, но и их представления о ситуации могут быть различны.

В этом случае агенты могут по-разному предвидеть результаты одного и того же совместного действия. Это порождает определенную специфику их взаимодействия. В частности, в игре двух агентов с одним общим целевым фактором и с разными желаемыми значениями для него становится возможным такое совместное воздействие агентов, что каждый из агентов в своих представлениях полностью достигает своей цели. Иначе говоря, несогласованные представления могут в некоторых случаях компенсировать существенное отличие в целевых установках и послужить причиной к полному согласию там, где его не могло бы быть в случае согласованных представлений.

## 1. УПРАВЛЕНИЕ СИТУАЦИЕЙ В МОДЕЛИ ЛИНЕЙНОЙ КОГНИТИВНОЙ КАРТЫ

Под *линейной когнитивной картой* понимается взвешенный орграф, вершины и дуги которого удовлетворяют описанным далее условиям, и задано правило изменения значений вершин (см. далее). Каждая вершина этого орграфа представляет *фактор* из предметной области. Множество всех

факторов обозначим  $M = \{1, \dots, m\}$ . Ориентированные дуги представляют *причинно-следственные связи* между факторами в рассматриваемой предметной области. Фактор-причина — это фактор, из которого выходит дуга, а фактор-следствие — это фактор, в который она входит. Далее, под *матрицей смежности орграфа*  $W$  будем понимать матрицу, элементы которой  $w_{ji} \in R$  соответствуют весам дуг, задающих силу и вид причинно-следственных связей. Абсолютное значение веса дуги  $|w_{ji}|$  задает силу причинно-следственной связи  $j$ -го фактора-причины на  $i$ -й фактор-следствие. Знак веса дуги задает вид связи: если  $w_{ji} > 0$ , то причинно-следственная связь  $j$ -го фактора на  $i$ -й фактор положительная, если  $w_{ji} < 0$ , то причинно-следственная связь отрицательная [1].

Пусть время дискретно и начальному состоянию системы соответствует нулевой момент времени. *Автономный импульсный процесс* во взвешенном орграфе определяется по правилу (1) с вектором начальных значений факторов  $x(0) = (x_1(0), x_2(0), \dots, x_m(0))$ ,  $x(0) \in R^m$ , и вектором  $p(0) = (p_1(0), p_2(0), \dots, p_m(0))$ ,  $p(0) \in R^m$ , задающим внешний импульс, вводимый в каждую вершину в нулевой момент времени [5].

$$x_i(t+1) = x_i(t) + p_i(t),$$

$$\text{где } p_i(t) = \begin{cases} p_i(0), & t = 0 \\ \sum_{j \in M} w_{ji} p_j(t-1), & t = 1, 2, 3, \dots, \end{cases} \quad (1)$$

где  $w_{ji}$  — элементы матрицы смежности  $W$  орграфа.

Зафиксируем дискретный момент времени  $T > 0$ . Вектор значений всех факторов  $x(T)$  определяется из выражения

$$\begin{aligned} x(T) &= x(0) + p(0) + p(1) + \dots + p(T-1) = \\ &= x(0) + p(0) + p(0)W + \dots + p(0)W^{T-1}. \end{aligned}$$

Если вынести за скобки общий множитель  $p(0)$ , то получим

$$\begin{aligned} x(T) &= x(0) + p(0)(E + W + \dots + W^{T-1}) = \\ &= x(0) + p(0) {}_T Q, \end{aligned}$$

где  $E$  — единичная матрица. Матрицу  ${}_T Q = E + W + \dots + W^{T-1}$  будем называть *матрицей достижимости воздействий к моменту времени  $T$*  для матрицы смежности  $W$ . Тогда сумма последовательных приращений для фактора  $x_j$  представима в виде

$$\sum_{t=0}^{T-1} p_j(t) = \sum_{k \in M} {}_T q_{kj} p_k(0), \quad (2)$$

где  ${}_T q_{kj}$  — элементы матрицы  ${}_T Q$ . Рассмотрим задачу управления слабоструктурированной ситуацией в модели линейной когнитивной карты. Под *управляющими воздействиями* будем понимать внешние импульсы, вводимые в каждую вершину в начальный момент времени  $p(0)$ ; причем  $p_j(0) = 0$ , если на  $j$ -ю вершину воздействие не оказывается. Под *результатом управления* понимается совокупность значений всех факторов в момент времени  $T$ :

$$x_j(T) = x_j(0) + \sum_{t=0}^T p_j(t), \quad j \in M. \quad (3)$$

Определение *цели управления* задается условиями, налагаемыми на результат управления  $x(T) = (x_1(T), x_2(T), \dots, x_m(T))$ ,  $x(T) \in R^m$ .

## 2. ОПИСАНИЕ ТЕОРЕТИКО-ИГРОВОЙ МОДЕЛИ КОНФЛИКТА

Рассмотрим задачу управления слабоструктурированной ситуацией, в которой управляющие воздействия в нулевой момент времени осуществляют сразу несколько агентов. Выбор управляющего воздействия агентом определяется *правилом индивидуального рационального выбора* [7]. При таком управлении возникает конфликтная ситуация или *игра* между агентами, т. е. взаимодействие агентов, в котором полезность каждого агента зависит как от его собственного действия (стратегии), так и от действий других агентов. Обозначим множество всех агентов через  $N = \{1, \dots, n\}$ . Каждый из агентов  $i \in N$  формально может быть представлен тройкой параметров  $\langle S_i, f_i, C_i \rangle$ :

— возможностями по оказанию управляющего воздействия на ситуацию, которые задаются *множеством стратегий*  $S_i$ ;

— целью управления, которая описывается в виде *функции полезности* для агента  $f_i$ ;

— знаниями агента об управляемой ситуации, которые задаются в виде *линейной когнитивной карты*  $C_i$ .

Рассмотрим каждый из параметров более подробно. Будем считать, что *линейные когнитивные карты* всех агентов  $C_1, C_2, \dots, C_n$  имеют одинаково упорядоченное множество факторов  $M = \{1, \dots, m\}$ , но могут отличаться причинно-следственными связями между факторами. Таким образом, когнитивные карты разных агентов имеют в общем виде различные матрицы смежности орграфов  $W^{(1)}, W^{(2)}, \dots, W^{(n)}$ . Однако вполне возможно, что у некоторых (в частном случае — у всех) агентов

матрицы смежности могут совпадать  $W^{(i)} = W^{(j)}$ . Будем считать, что дискретное время в моделях когнитивных карт у всех агентов протекает одинаково. Начальный момент времени фиксирован для всех и равен 0. Всем агентам известен вектор начальных значений факторов  $x(0) \in R^m$ .

Агент с номером  $i \in N$  располагает непустым подмножеством факторов  $M_i \in M$ , доступных ему для управляющего воздействия. Будем называть  $M_i$  *множеством управляемых факторов*  $i$ -го агента. Для любых двух агентов  $i, j \in N$ :  $M_i \cap M_j = \emptyset$ , и  $\cup_{k \in N} M_k \subseteq M$ . Число факторов в множестве  $M_i$  будем обозначать  $m_i$ .

В векторе совместных управляющих воздействий  $p(0) = (p_1(0), p_2(0), \dots, p_m(0))$  отражается воздействие каждого из агентов. *Стратегией  $i$ -го агента*  $s_i$  будем считать вектор, состоящий из упорядоченных компонентов вектора  $p(0)$  с номерами из множества  $\{k_1, k_2, \dots, k_{m_i}\} = M_i$ :  $s_i = (p_{k_1}(0), p_{k_2}(0), \dots, p_{k_{m_i}}(0))$ . Каждый агент воздействует на ситуацию, задавая значения лишь для «своих» компонентов вектора  $p(0)$ . Остальные компоненты вектора  $p(0)$ , на которые не воздействует ни один из агентов, равны нулю:  $(\forall j \in M \setminus \cup_{k \in N} M_k), p_j(0) = 0$ .

Воздействие требует затрат определенного количества ресурсов, которые, как правило, ограничены. Введем простейшие ограничения на управляющие воздействия по каждому фактору в виде отрезка допустимых значений:  $(\forall j \in \cup_{k \in N} M_k)$

$p_j(0) \in [p_j^{\min}, p_j^{\max}]$ , где  $p_j^{\min}, p_j^{\max} \in R$ . Тогда *множество стратегий  $i$ -го агента*  $S_i$  можно представить как декартово произведение  $m_i$  отрезков  $[p_{k_1}^{\min}, p_{k_1}^{\max}] \times [p_{k_2}^{\min}, p_{k_2}^{\max}] \times \dots \times [p_{k_{m_i}}^{\min}, p_{k_{m_i}}^{\max}]$ . Фактически множество стратегий  $i$ -го агента  $S_i$  является  $m_i$ -мерным прямоугольным параллелепипедом, однако далее для краткости будем называть его  *$m_i$ -мерным гиперкубом*, а множество всех возможных стратегий агентов  $S_1 \times \dots \times S_n$  — просто *гиперкубом*. Будем рассматривать  $\{s_i\}_{i \in N}$  как вектор стратегий всех агентов  $(s_1, s_2, \dots, s_n) \in S_1 \times \dots \times S_n$ .

Для каждого агента определим на множестве результатов управления функцию полезности  $f_i(x_1(T), x_2(T), \dots, x_m(T))$ . *Цель управления  $i$ -го агента* заключается в максимизации значения функции  $f_i$ . Рассмотрим по отдельности два случая, для которых будут построены функции полезности.

1. Агенту  $i$  важно неограниченное увеличение (либо уменьшение) значения фактора  $x_j$ . Тогда в



случае увеличения значения фактора необходимо максимизировать выражение  $(x_j(T) - x_j(0))$ , а в случае уменьшения значения фактора необходимо максимизировать выражение  $-(x_j(T) - x_j(0))$ .

2. Агенту  $i$  важно неограниченное приближение значения фактора  $x_j$  к некоторому выгодному для него значению  $x_{ij}^*$ . Тогда будем считать, что для него желательно максимизировать выражение  $-(x_j(T) - x_{ij}^*)^2$ .

Отметим, что выражения в п. 1 и 2 записаны так, чтобы максимум их значений соответствовал цели агента. В случае, когда агенту «небезразличны» значения более одного фактора, можно записать свертку по вышеописанным выражениям для таких факторов с соответствующими коэффициентами, каждый из которых интерпретируется как «доля важности» ограничений на значение соответствующего фактора. Для обоих случаев запишем соответствующий вид функции полезности для произвольного  $i$ -го агента.

Функция полезности для случая 1 имеет вид:

$$f_i(x_1(T), x_2(T), \dots, x_m(T)) = \sum_{j \in M} \gamma_{ij}(x_j(T) - x_j(0)), \quad (4)$$

где  $|\gamma_{ij}|$  — «доля важности» значения  $j$ -го фактора среди остальных факторов для  $i$ -го агента,  $\gamma_{ij} \in [-1, 1]$ , сумма величин  $|\gamma_{ij}|$  равна единице для фиксированного  $i \in N$ . Знак коэффициента  $\gamma_{ij}$  отражает направление изменения значения фактора, выгодное для агента. Если  $\gamma_{ij} > 0$ , то  $i$ -й агент стремится неограниченно увеличить значение  $j$ -го фактора. Если  $\gamma_{ij} < 0$ , то  $i$ -й агент стремится неограниченно уменьшить значение  $j$ -го фактора. Если  $\gamma_{ij} = 0$ , то  $i$ -му агенту безразлично значение  $j$ -го фактора.

Функция полезности для второго случая будет иметь вид:

$$f_i(x_1(T), x_2(T), \dots, x_m(T)) = - \sum_{j \in M} \gamma_{ij}(x_j(T) - x_{ij}^*)^2, \quad (5)$$

где  $\gamma_{ij}$  — «доля важности» значения  $j$ -го фактора для  $i$ -го агента,  $\gamma_{ij} \in [0, 1]$ , сумма величин  $\gamma_{ij}$  равна единице для фиксированного  $i \in N$ .

Фактор, в выражении для которого в записи функции полезности для  $i$ -го агента в виде (4) или (5) коэффициент  $\gamma_{ij} \neq 0$ , будем называть *целевым фактором* для  $i$ -го агента.

Определив все параметры игры, запишем ее в нормальной форме:

$$\Gamma_C = \{N, \{S_i\}_{i \in N}, \{f_i\}_{i \in N}, \{C_i\}_{i \in N}\}. \quad (6)$$

Перечислим в явном виде допущения, из которых мы будем исходить при рассмотрении игры (6):

- 1) множества факторов в когнитивных картах всех агентов совпадают;
- 2) множества управляемых факторов для всех агентов заданы и не пересекаются;
- 3) агенты знают когнитивные карты, функции полезности и множества стратегий друг друга;
- 4) каждый агент верит в адекватность лишь собственной когнитивной карты;
- 5) п. 1—4 являются общим знанием среди агентов.

Далее мы будем рассматривать две игры, в каждой из которых функции полезности одновременно всех агентов описываются либо в виде (4), либо в виде (5). Для краткости, пользуясь терминологией из работы [3], будем их называть *моделью с нефиксированной целью управления* и *моделью с фиксированной целью управления* соответственно. Решения обеих игр будем искать в соответствии с концепцией равновесия Нэша в чистых стратегиях [7]. В зависимости от игры будут определяться *равновесие в доминантных стратегиях* в случае модели с нефиксированной целью управления, и *равновесие Нэша в чистых стратегиях* в модели с фиксированной целью управления.

### 3. ПОИСК РЕШЕНИЯ ДЛЯ МОДЕЛИ С НЕФИКСИРОВАННОЙ ЦЕЛЬЮ УПРАВЛЕНИЯ

Рассмотрим игру (6), в которой все функции полезности агентов записаны в виде (4). Отметим, что подобная задача решена в статье [6] и переформулирована в используемых нами обозначениях в работе [8]. В этих работах рассмотрен случай, когда у всех агентов когнитивные карты совпадают или, иначе, они все играют на одной когнитивной карте. В настоящей статье постановка задачи другая: когнитивные карты агентов различны. Однако рассуждения, изложенные в работах [6, 8], могут быть успешно повторены и для нашей постановки.

Выберем произвольного агента  $i \in N$ , и далее будем рассматривать ситуацию с его точки зрения. Преобразуем его функцию полезности  $f_i(x_1(T), x_2(T), \dots, x_m(T))$  к виду, в котором явно отражена зависимость его выигрыша от действий всех агентов  $g_i(s_1, s_2, \dots, s_n)$ . Для этого в выражение (4) вместо  $x_j(T)$  подставим правую часть выражения (3). Получим:

$$\begin{aligned} \sum_{j \in M} \gamma_{ij}(x_j(T) - x_j(0)) &= \sum_{j \in M} \gamma_{ij}(x_j(0) + \sum_{t=0}^T p_j(t) - \\ &- x_j(0)) = \sum_{j \in M} \gamma_{ij} \left( \sum_{t=0}^T p_j(t) \right). \end{aligned}$$

Далее, используя выражение (2), получаем

$$\begin{aligned} \sum_{j \in M} \gamma_{ij} \left( \sum_{t=0}^T p_j(t) \right) &= \sum_{j \in M} \gamma_{ij} \left( \sum_{k \in M} T q_{kj}^{(i)} p_k(0) \right) = \\ &= \sum_{j \in M} \sum_{k \in M} \gamma_{ij} T q_{kj}^{(i)} p_k(0) = \sum_{k \in M} \left[ \left( \sum_{j \in M} \gamma_{ij} T q_{kj}^{(i)} \right) p_k(0) \right]. \end{aligned}$$

Заметим, что при подстановке выражения (2) использовались элементы  $T q_{kj}^{(i)}$  матрицы достижимости воздействий к моменту времени  $T$   $T Q^{(i)}$  для матрицы смежности орграфа  $W^{(i)}$ , т. е. знания  $i$ -го агента о ситуации. Это обосновано четвертым допущением, сделанным в предыдущем параграфе, о том, что каждый агент верит в адекватность лишь собственной когнитивной карты.

Обозначим  $T \alpha_{ik}^{(i)} = \left( \sum_{j \in M} \gamma_{ij} T q_{kj}^{(i)} \right)$ . Тогда функцию полезности  $i$ -го агента можно записать в виде, называемом *целевой функцией  $i$ -го агента*:

$$\begin{aligned} g_i &= \sum_{k \in M} T \alpha_{ik}^{(i)} p_k(0) = \\ &= \sum_{k \in M_i} T \alpha_{ik}^{(i)} p_k(0) + \sum_{k \in M \setminus M_i} T \alpha_{ik}^{(i)} p_k(0). \end{aligned} \quad (7)$$

Представление целевой функции  $g_i$  в виде (7) позволяет аддитивно выделить зависимость ее значения от выбранной  $i$ -м агентом стратегии. Согласно лемме из работы [7] у агента  $i$  есть доминантная стратегия, следовательно, *решением в модели с нефиксированной целью управления будет совокупность доминантных стратегий агентов  $s_i^*$  (равновесие в доминантных стратегиях). Доминантные стратегии агентов вычисляются по формуле*

$$p_j^*(0) = \begin{cases} p_j^{\min}, & \text{sign}(T \alpha_{ij}^{(i)}) < 0, \\ p_j^{\max}, & \text{sign}(T \alpha_{ij}^{(i)}) > 0, \quad \forall j, j \in M_i, \\ 0, & \text{для определенности, иначе.} \end{cases}$$

#### 4. ПОИСК РЕШЕНИЯ ДЛЯ МОДЕЛИ С ФИКСИРОВАННОЙ ЦЕЛЬЮ УПРАВЛЕНИЯ

Рассмотрим игру (6), в которой все функции полезности агентов записаны в виде (5). Вновь выберем произвольного агента  $i \in N$ , и далее будем рассматривать ситуацию с его точки зрения. Преобразуем его функцию полезности  $f_i(x_1(T), x_2(T), \dots, x_m(T))$  к виду, в котором явно отражена зависимость его выигрыша от действий всех аген-

тов  $g_i(s_1, s_2, \dots, s_n)$ . Для этого в выражение (5) вместо  $x_j(T)$  подставим правые части выражений (3) и (2):

$$\begin{aligned} \sum_{j \in M} \gamma_{ij} [-(x_j(T) - x_j^*)^2] &= \\ &= - \sum_{j \in M} \gamma_{ij} (x_j(0) + \sum_{k \in M} T q_{kj}^{(i)} p_k(0) - x_j^*)^2. \end{aligned}$$

Заметим, что при подстановке выражения (2), как и в предыдущем разделе, при записи функции использовались элементы  $T q_{kj}^{(i)}$  матрицы достижимости воздействий к моменту времени  $T$   $T Q^{(i)}$  для матрицы смежности орграфа  $W^{(i)}$ . Как и при записи функции (7), это обосновано четвертым допущением о том, что каждый агент верит в адекватность лишь собственной карты. Таким образом, из функции полезности (5) получаем целевую функцию

$$\begin{aligned} g_i &= - \sum_{j \in M} \gamma_{ij} \left( c_{ij} + \sum_{k \in M} T q_{kj}^{(i)} p_k(0) \right)^2 = \\ &= - \sum_{j \in M} \gamma_{ij} (c_{ij} + T q_{1j}^{(i)} s'_i + T q_{2j}^{(i)} s'_{-i})^2. \end{aligned} \quad (8)$$

Здесь  $c_{ij} = x_j(0) - x_j^*$ ;  $T q_{1j}^{(i)}$  и  $T q_{2j}^{(i)}$  — векторы, составленные из соответствующих коэффициентов  $T q_{kj}^{(i)}$  в первой и второй суммах соответственно;  $s'_i$  и  $s'_{-i}$  — транспонированные вектор стратегий и вектор обстановки игры соответственно для  $i$ -го агента. Функция (8) строго вогнута по переменной-стратегии  $s_i$ ;  $\forall \alpha \in (0, 1) g_i(\alpha s_i^* + (1 - \alpha) s_i^{**}, s_{-i}) > \alpha g_i(s_i^*, s_{-i}) + (1 - \alpha) g_i(s_i^{**}, s_{-i})$ . Кроме того, функция (8) непрерывна по всем своим переменным, как суперпозиция непрерывных функций. Согласно соответствующей теореме из работы [7] для модели с фиксированной целью управления существует равновесие Нэша в чистых стратегиях.

Отметим, что модель с фиксированной целью управления формально похожа на олигополию Курно [9]. Далее для поиска решения в модели с фиксированной целью управления, аналогично решению задачи олигополии Курно, построим систему уравнений для поиска наилучшего ответа

$$\text{каждого из агентов: } \frac{\partial g_i(s_i^*, s_{-i}^*)}{\partial p_l(0)} = 0, \quad \forall l \in M_i, \quad \forall i \in N.$$

Запишем эту систему уравнений в явном виде после соответствующего преобразования:

$$\begin{aligned} \sum_{k \in M} \left( \sum_{j \in M} \gamma_{ij} T q_{lj}^{(i)} T q_{kj}^{(i)} \right) p_k(0) &= - \sum_{j \in M} \gamma_{ij} T q_{lj}^{(i)} c_{ij}, \\ \forall l \in M_i, \quad \forall i \in N. \end{aligned} \quad (9)$$



Как отмечалось ранее, в выражении (9)  $p_k(0) = 0$ ,  $\forall k \notin \cup_{i \in N} M_i$ , т. е. по неуправляемым факторам управляющие воздействия отсутствуют. Подобно решению задачи олигополии Курно [9] можно утверждать, что *если решение системы уравнений (9) существует и принадлежит гиперкубу  $S_1 \times \dots \times S_n$ , то оно является равновесием Нэша в чистых стратегиях для модели с фиксированной целью управления*. Также можно показать, что *если точка равновесия Нэша в чистых стратегиях для модели с фиксированной целью управления принадлежит внутренности гиперкуба  $S_1 \times \dots \times S_n$ , то она является решением системы уравнений (9)*.

Из этих двух утверждений и из существования равновесия Нэша в чистых стратегиях для модели с фиксированной целью управления следует справедливость того, что *если ни одно из решений системы уравнений (9) не принадлежит гиперкубу  $S_1 \times \dots \times S_n$ , то точка равновесия Нэша в чистых стратегиях для модели с фиксированной целью управления лежит на границе гиперкуба  $S_1 \times \dots \times S_n$* .

## 5. ПРИМЕР

Рассмотрим пример взаимодействия двух агентов, когнитивные карты которых содержат три фактора  $A$ ,  $B$  и  $C$  (рис. 2).

Начальное значение каждого фактора указано в вершинах графа:  $A = 1$ ,  $B = -7$ ,  $C = 11$ . Пусть  $A$  — целевой фактор для обоих агентов,  $B$  — управляемый фактор первого агента,  $C$  — управляемый фактор второго агента. За момент времени для регистрации результата управления примем  $T = 3$ . Управляющие воздействия на фактор  $B$  ограничены отрезком  $[-10, 10]$ , а на фактор  $C$  — отрезком  $[-15, 15]$ . Эти отрезки являются множеством стратегий первого и второго агентов соответственно:  $S_1 = [-10, 10]$ ,  $S_2 = [-15, 15]$ . Рассмотрим случай с нефиксированными целями управления, когда цели агентов противоположны. Пусть целью пер-

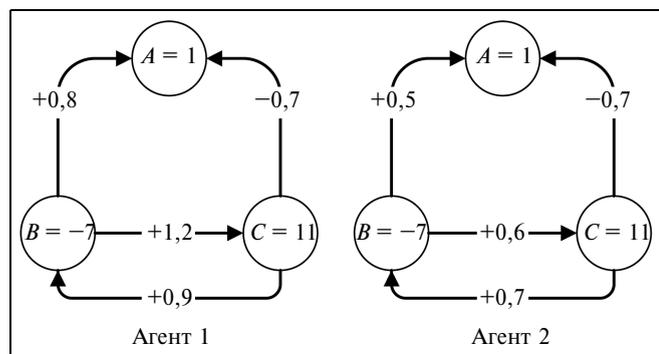


Рис. 2. Когнитивные карты двух агентов

вого агента будет неограниченное уменьшение значения фактора  $A$ , а целью второго — его неограниченное увеличение. В этом случае целевые функции агентов примут вид:  $f_1 = -(x_A(3) - 1) = 0,04p_B(0) - 0,02p_C(0)$ ,  $f_2 = (x_A(3) - 1) = 0,08p_B(0) - 0,35p_C(0)$ . Очевидно, решением игры будет равновесие в доминантных стратегиях  $(p_B^*(0), p_C^*(0)) = (10, -15)$ . Это решение — единственное оптимальное по Парето в игре, так как доставляет максимальный возможный выигрыш и первому, и второму агенту одновременно, и другого такого набора стратегий нет. Этот пример иллюстрирует возможность того, что агенты с противоположными исходными целями могут образовать коалицию «по действиям» из-за несогласованности представлений. Можно видеть, что, если в данном примере цели агентов были бы одинаковы, например, оба агента стремились бы к неограниченному уменьшению значения фактора  $A$ , то действия одного агента мешали бы другому в достижении поставленной цели, и наоборот. Отмеченный факт может служить иллюстрацией того, что агенты с одинаковыми целями, но разными представлениями о ситуации, могут вступить в конфронтацию. И наоборот — агенты с противоположными целями и различными представлениями о ситуации могут действовать совместно.

Рассмотрим тот же пример взаимодействия двух агентов в случае с фиксированными целями управления. Пусть первому агенту важно неограниченное приближение значения фактора  $A$  к 0, а второму агенту — к 2. В этом случае целевые функции агентов примут вид:  $f_1 = -(x_A(3) - 0)^2 = -(1 - 0,04p_B(0) + 0,02p_C(0))^2$ ,  $f_2 = -(x_A(3) - 2)^2 = -(0,08p_B(0) - 0,35p_C(0) - 1)^2$ . Решением системы уравнений (9) для данной задачи будет вектор  $(26,61; 3,23)$ , который, очевидно, не принадлежит множеству всех возможных наборов стратегий агентов  $S_1 \times S_2$ . В этом случае единственным равновесием Нэша будет точка на границе множества  $S_1 \times S_2$ :  $(10; -0,57)$ . Значения целевых функций у агентов:  $f_1(10; -0,57) = -0,35$ ,  $f_2(10; -0,57) = 0$ . Однако первый агент может улучшить свой выигрыш, если сообщит заведомо ложную информацию о своей когнитивной карте второму агенту, искажая силу влияния  $B \rightarrow C$  с  $+1,2$  на  $+0,5$ , и  $C \rightarrow B$  с  $+0,9$  на  $+0,3$ . Отметим, что этот пример выпадает из наших предположений, сделанных в § 2 о том, что все агенты знают истинные когнитивные карты друг друга.

В этом случае равновесие Нэша для второго агента смещается в точку  $(-6,71; -4,39)$ . Этим может воспользоваться первый агент, сохраняя свою истинную наилучшую стратегию. Построим *граф рефлексивной игры* (рис. 3) и рассчитаем *информационное равновесие* [10].

Вершины графа соответствуют реальным (1 и 2) и фантомным (21) агентам, т. е. попарно нетождественным структурам информированности. Дуги графа отражают взаимную информированность агентов: направление стрелки от одного агента к другому обозначает, что второй агент адекватно информирован о первом. Решением будет *информационное равновесие*  $(10; -4,39)$

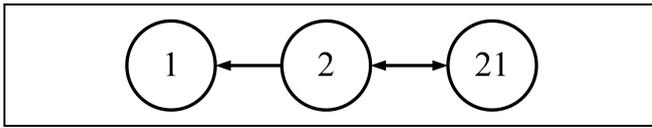


Рис. 3. Граф рефлексивной игры

[10], значения целевых функций агентов будут следующими:  $f_1(10; -4,39) = -0,26$ ;  $f_2(10; -4,39) = -1,79$ . При сравнении выигрышей первого агента в двух случаях видно, что он смог добиться его увеличения от  $-0,35$  до  $-0,26$  путем дезинформации второго агента.

### ЗАКЛЮЧЕНИЕ

Рассмотрены модели конфликта интересов, в которых агенты представлены различными описаниями собственных знаний о ситуации в целом, формализованными в виде линейных когнитивных карт. Получены решения для обеих моделей в рамках концепции решения равновесия Нэша в чистых стратегиях. Эти результаты опираются на сделанные в § 2 допущения. Еще раз отметим два важных момента.

- Формальному описанию подвергается внутреннее видение агентами законов развития ситуации во времени. В этом случае агенты по-разному предвидят результаты одного и того же действия, но каждый агент верит в адекватность лишь собственного результата. Как следствие, в приведенных моделях выигрышем агента считается величина, рассчитанная, исходя из представлений лишь этого агента о своем выигрыше: агент, принимая решение, понимает не только то, что у остальных агентов другие цели, но и то, что они видят мир иначе. Такая постановка приводит к целесообразности рассмотрения рефлексивных игр [9, 10] на когнитивных картах.
- В рассмотренных моделях знания агентов о ситуации представлены разными когнитивными картами и отсутствует какая-либо информация о том, какая из когнитивных карт адекватно описывает ситуацию, если такой картой обладает кто-либо из агентов. Иначе говоря, решаемая нами задача прогноза действий агентов не зависит от того, кто из них (а, может быть, и никто) адекватно представляет ситуацию в целом.

В рамках примеров (см. § 5) были продемонстрированы некоторые эффекты, возникающие при моделировании взаимодействий агентов с несогласованными представлениями. Например, то, что агенты с противоположными исходными целями и

различной информированностью могут образовать коалицию «по действиям». В частности, в первом примере именно различие когнитивных карт у агентов послужило причиной максимизации их общего благосостояния. Также представлен случай, когда одинаковые цели и разные представления о ситуации приводят к конфронтации участников. В примере с фиксированными целями показана возможность увеличения выигрыша агента путем дезинформации им своих оппонентов относительно собственной когнитивной карты. Все это позволяет ставить и решать задачи информационного управления на когнитивных картах, причем управление может осуществляться как самими агентами, так и внешними по отношению к ним сторонами для достижения собственных целей.

### ЛИТЕРАТУРА

1. Кузнецов О.П., Кулинич А.А., Марковский А.В. Анализ влияния при управлении слабоструктурированными ситуациями на основе когнитивных карт // Человеческий фактор в управлении. — М.: КомКнига, 2006. — С. 313–344.
2. Кузнецов О.П. Интеллектуализация поддержки управляющих решений и создание интеллектуальных систем // Проблемы управления. — 2009. — № 3.1. — С. 64–72.
3. Максимов В.И., Корноушенко Е.К. Аналитические основы применения когнитивного подхода при решении слабоструктурированных задач // Труды ИПУ РАН. — М., 1999. — Т. II. — С. 95–109.
4. Максимов В.И. Структурно-целевой анализ развития социально-экономических ситуаций // Проблемы управления. — 2005. — № 3. — С. 30–38.
5. Робертс Ф. Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. — М.: Наука, 1986. — 325 с.
6. Новиков Д.А. «Когнитивные игры»: линейная импульсная модель // Проблемы управления. — 2008. — № 3 — С. 14–22.
7. Губко М.В., Новиков Д.А. Теория игр в управлении организационными системами. — М.: СИНТЕГ, 2005. — 138 с.
8. Куливец С.Г. Теоретико-игровые модели на линейных когнитивных картах // Сб. науч. тр. V междунар. науч.-техн. конф. «Интегрированные модели и мягкие вычисления в искусственном интеллекте». — М., 2009. — С. 379–386.
9. Чхартшвили А.Г. Теоретико-игровые модели информационного управления. — М.: ЗАО «ПМСОФТ», 2004. — 227 с.
10. Новиков Д.А., Чхартшвили А.Г. Рефлексивные игры. — М.: СИНТЕГ, 2003. — 160 с.

Статья представлена к публикации членом редколлегии чл.-корр. РАН Д.А. Новиковым.

Куливец Сергей Геннадьевич — аспирант, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ☎ (495) 334-76-39, ✉ skulivec@yandex.ru.

# МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОПТИМИЗАЦИИ СТРУКТУРЫ ИЕРАРХИЧЕСКОГО МЕНЮ

М.В. Губко, А.И. Даниленко

Предложена математическая модель оптимизации структуры иерархических меню и каталогов. Каждому элементу меню поставлена в соответствие его популярность и решена задача дискретной оптимизации — выбора структуры меню, минимизирующей среднее время поиска. Доказано, что все панели оптимального меню должны предлагать пользователю одинаковое число альтернатив с одинаковым соотношением популярностей. Отмечено, что модель позволяет сравнивать типы меню и выбрать из них наилучший. Предложен алгоритм построения оптимального меню, учитывающий как семантические ограничения, так и результаты оптимизации. Алгоритм проиллюстрирован примером оптимизации меню мобильного телефона.

**Ключевые слова:** иерархические меню, оптимизация иерархии, автоматизация построения меню, меню мобильного устройства.

## ВВЕДЕНИЕ

Иерархические меню являются популярным методом доступа к командам и данным. Примерами могут служить командные меню компьютерных программ (рис. 1), каталоги Интернет-сайтов (рис. 2), меню мобильных телефонов, интерактивные телефонные системы и т. п.

Меню обычно является если не единственным, то самым полным каталогом команд или ссылок в системе, поэтому удобство меню является важным критерием качества интерфейса в целом. Внешний вид и пользовательские качества меню могут отличаться в зависимости от места его размещения и задач, которые оно решает. Например, меню мобильных телефонов проектируется с учетом ограниченного размера экрана устройства, у голосовых меню нет визуального интерфейса, что вносит дополнительные ограничения на их структуру и содержание. Меню компьютерных программ (см. рис. 1) должны иметь стандартный вид, знакомый пользователям операционной системы. В дизайне Интернет-сайтов меньше правил, поэтому при проектировании меню сайтов и каталогов открывается большой простор для творчества. Например, для больших каталогов Интернет-сайтов часто применяется двухуровневая структура меню (рис. 2) в которой на одной панели меню отобра-

жаются как сами категории, так и подчиненные им подкатегории.

Эффективность меню, как правило, определяется средним временем доступа к искомому элементу. При этом для конкретного набора элементов существует множество способов их группировки в иерархию категорий. Задача дизайна меню состоит в поиске иерархии категорий, обеспечивающей минимальное среднее время доступа с учетом выбора для каждой панели наиболее подходящего вида меню.

## 1. ОБЗОР ЛИТЕРАТУРЫ

Еще до развития компьютерных систем схожая проблематика рассматривалась в теории вопросов [1]. Помимо подробного описания теории в работе [1] проведена аналогия с теорией оптимизации организационных иерархий, использующей тот же математический аппарат, что и настоящая статья. Вследствие широкого распространения графических компьютерных интерфейсов в 1980-х гг. началась активная разработка формальных методов оптимизации структуры меню (содержательный, хотя и несколько устаревший, обзор приведен в работе [2]). В настоящее время больше внимания уделяется улучшению пользовательских качеств Интернет-сайтов, мобильных устройств и бытовой электроники [3—5]. Ряд публикаций пос-

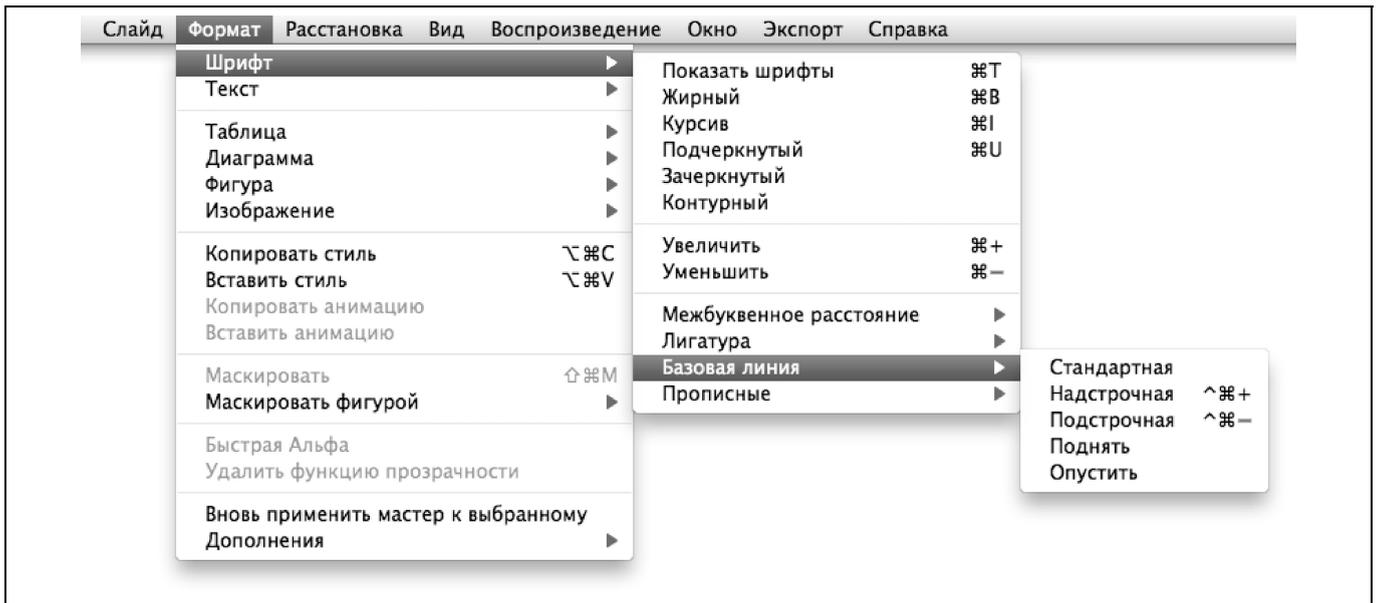


Рис. 1. Фрагмент командного меню программы «Apple Keynote»

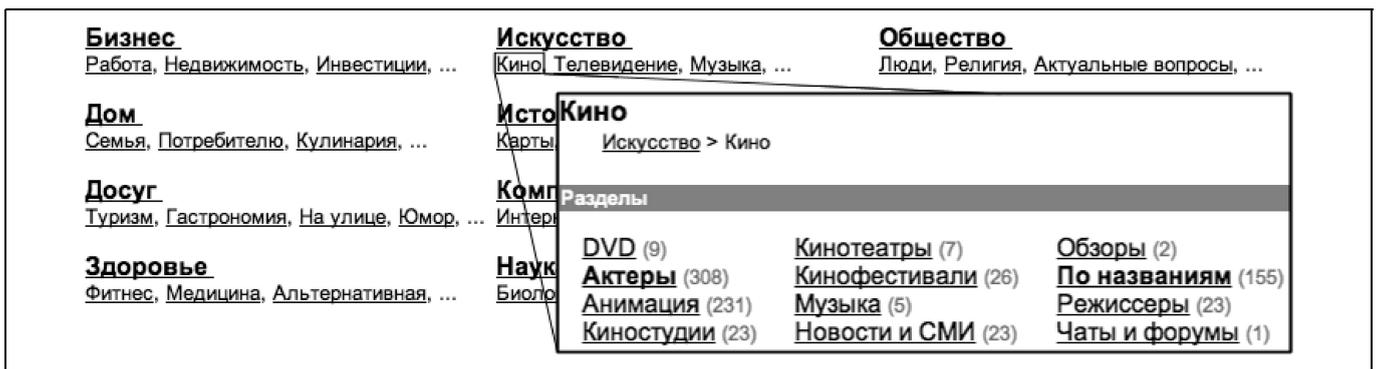


Рис. 2. Фрагмент двухуровневого меню Интернет-каталога Google Directory

вещен изучению различных принципов организации меню. Пункты меню могут быть расположены по алфавиту [6], в случайном порядке [7, 8], или упорядочены по принадлежности к подкатегориям [9, 10]. Структура и отображение пунктов меню может изменяться в зависимости от частоты их использования [11, 12].

Отдельное направление исследований — влияние ширины и глубины иерархии на время поиска. Главный вопрос заключается в том, сколько альтернатив следует предлагать пользователю на одной панели меню (это число называется шириной иерархии) и сколько при этом должно быть уровней в иерархии (другими словами, какая глубина иерархии). Существуют различные подходы к этому вопросу. Классическая модель [8] (позже расширенная в работе [10]) рассматривает построение симметричного меню постоянной ширины над

элементами с одинаковой популярностью в предположении, что выбор любого пункта меню требует одинакового времени. Результатом предложенной линейной модели стало выражение для оптимальной ширины иерархии (в среднем 6—8 пунктов). Полученные результаты далеки от экспериментальных исследований, и в работе [6] была предложена альтернативная модель с подобными ограничениями, но с логарифмической зависимостью времени поиска от ширины иерархии, предсказывающая оптимальность более широких иерархий (см. также [11]). Настоящую статью можно рассматривать как продолжение этих работ, так как предложенная модель обобщает подходы, описанные в работах [6, 8, 10], снимая ряд ограничений.

Важную роль в исследовании структур меню играют многочисленные экспериментальные исследования [7, 9, 13—17], базирующиеся на раз-



личных моделях и предпосылках. Несмотря на выявленные в ходе исследований общие принципы, различие подходов и искусственность экспериментов не позволили прийти к единому мнению о влиянии ширины и глубины иерархии на удобство меню.

Несколько работ затрагивают вопрос переменной ширины иерархии в меню. Так, в работах [17, 18] показывается, что более широкие на последних уровнях меню эффективнее, чем меню, сужающиеся к концу. Однако, например, в работах [16, 19], существенных различий между такими структурами выявлено не было.

Существует также направление исследований, посвященное изучению особенностей специфических типов меню: голосовых меню [20], меню мобильных телефонов [3, 5] и др.

Таким образом, существующие исследования структуры иерархических меню существенно различаются как в предположениях о свойствах системы и поведении пользователей, так и в применяемых методах. Кроме того, полученные результаты, как теоретические, так и экспериментальные, сложно применять на практике. Теоретические результаты получены при сильных ограничениях, не применимых к реальным меню. Экспериментальные методы исследуют отдельно взятую область, и их результаты не всегда возможно обобщить. Такая ситуация привела к появлению работ [21], в которых сравниваются различные методы и определяется их применимость в зависимости от условий и стратегий пользователей. Но до сих пор не существует единой методики, описывающей принципы построения меню для различных случаев. Такая методика должна содержать элементы оптимизации структуры меню и позволять сравнивать различные типы и структуры меню на предмет их удобства для пользователей. В настоящей статье исследуется математическая модель, отвечающая на поставленные вопросы с помощью результатов математической теории оптимизации иерархических структур [22, 23]. Общий подход, предложенный в рассматриваемой модели, позволяет ослабить ряд допущений существующих моделей и использовать полученные теоретические результаты для разработки средств автоматизации построения меню.

## 2. ПОСТАНОВКА ЗАДАЧИ

Иерархическое меню строится на основе множества элементов  $N = \{1, \dots, n\}$  (например, команд или ссылок в Интернет-каталоге), к которым необходимо получать доступ посредством меню. Через  $\mu(w)$  обозначим популярность элемента  $w \in N$  — вероятность того, что этот элемент потребуется пользователю.

При построении меню множество элементов  $N$  разбивается на категории, затем элементы каждой категории могут быть разбиты на подкатегории и т. д. Панелью меню для категории  $s \subseteq N$  будем называть совокупность самой категории  $s$  и набора содержащихся в ней вариантов (подкатегорий или терминальных элементов). Например, на рис. 1 раскрыты четыре панели меню. Категории верхнего уровня («Слайд», «Формат» и т. д.) расположены в главной панели меню, а категории следующих уровней принадлежат вложенным панелям. Такое многоуровневое разбиение элементов на категории формирует древовидную структуру меню. Листьями дерева служат элементы из множества  $N$ , остальные вершины соответствуют категориям.

Каждая категория однозначно характеризуется набором  $s \subseteq N$  входящих в нее элементов нижнего уровня. При этом вероятность того, что пользователь попадет в панель меню, соответствующую категории  $s$ , равна  $\mu_s := \sum_{w \in s} \mu(w)$ <sup>1</sup>.

Если пользователь попадает в панель меню категории  $s \subseteq N$ , содержащую  $k$  вариантов с популярностями  $\mu_1, \dots, \mu_k$ , то  $\mu_1 + \dots + \mu_k = \mu_s$ . Тогда  $y_i = \mu_i / \mu_s$  — условная вероятность выбора варианта  $i$  при нахождении в панели меню категории  $s$ . Другими словами, вектор  $y_1, \dots, y_k$  определяет пропорцию ( $y_1 + \dots + y_k = 1$  по определению), в соответствии с которой поделены популярности вариантов, содержащихся в категории  $s$ .

Будем пользоваться средним временем, затрачиваемым пользователем на поиск элемента (средним временем одной пользовательской сессии) как критерием качества меню. Как правило, считается [6—10, 14—17], что цель пользователя — один элемент нижнего уровня, например, одна запись в меню доступа к базам данных или одна команда в командных меню. Таким образом, если обозначить через  $t(w)$  время поиска пользователем элемента  $w \in N$ , то необходимо минимизировать среднее время доступа к одному элементу

$$T = \sum_{w \in N} \mu(w)t(w).$$

С другой стороны, поскольку все время пользователь проводит в поиске подходящих вариантов в категориях, среднее время сессии представимо в виде суммы средних времен пребывания в каждой панели меню. Время, проведенное в панели меню, зависит в первую очередь от набора предоставленных вариантов. Кроме этого, на время выбора существенно влияют пользовательские качества ме-

<sup>1</sup> Это верно, если пользователь не совершает ошибок при выборе вариантов в меню.

ню. Так, варианты меню могут быть отсортированы по алфавиту или по содержанию, пункты на панели меню могут быть расположены горизонтально, вертикально или в ячейках прямоугольной матрицы, категории представлены своими названиями, пиктограммами или и тем и другим, и т. д. Условимся называть типом меню совокупность всех факторов, влияющих на внешний вид<sup>2</sup> и пользовательские качества меню, за исключением числа и состава самих пунктов меню. Обозначим через  $\Omega$  множество всевозможных типов меню, применимых в рассматриваемом контексте.

Для конкретного типа меню  $\theta \in \Omega$  обозначим через  $t_i(k, \theta)$  среднее время, необходимое, чтобы найти и выбрать  $i$ -й вариант<sup>3</sup> в панели меню, содержащей  $k$  альтернатив. Это время включает в себя как время на восприятие и анализ вариантов, так и на совершение выбора (например, перемещение курсора или нажатие клавиши). Кроме набора вариантов и типа меню на время выбора влияют личные качества пользователей (квалификация, возраст), технические ограничения оборудования и ряд других факторов, которые неподвластны разработчику меню. Если информация о потенциальных пользователях, аппаратных задержках и других внешних параметрах доступна на этапе проектирования, ее следует включить в величину  $t_i(k, \theta)$ . Такая детализация позволяет получить более точные (но менее универсальные) рекомендации касательно структуры меню. При отсутствии подобной информации следует использовать средние значения<sup>4</sup>, которые позволяют получить более универсальные, но «оптимальные в среднем» параметры меню. Основным источником для определения времен  $t_i(k, \theta)$  служат экспериментальные исследования.

Если пользователь попал в панель меню категории  $s \subseteq N$  с  $k(s)$  подкатегориями  $s_1, \dots, s_k$ , то, умножив время выбора  $t_i(k, \theta)$  каждого варианта на соответствующую условную вероятность  $y_i(s) := \mu_{s_i} / \mu_s$  его выбора, получим среднее время, которое пользователь потратит на выбор варианта в этой панели меню:

$$t(y_1(s), \dots, y_{k(s)}(s), \theta) = \sum_{i=1}^{k(s)} y_i(s) t_i(k(s), \theta).$$

<sup>2</sup> Далее для удобства изложение ведется в «визуальных» терминах, но те же результаты верны и для других типов меню, в том числе голосовых.

<sup>3</sup> Считается, что для каждого сочетания  $k$  и  $\theta$  задана некоторая нумерация пунктов меню от 1 до  $k$ .

<sup>4</sup> Далее показано, что временные затраты входят в критерий оптимизации линейно, поэтому можно использовать усредненные параметры.

В общем случае для каждой панели может использоваться свой тип меню, который нужно выбрать в ходе оптимизации. Это можно делать независимо для каждой панели, выбирая тип меню, минимизирующий время нахождения в ней:

$$t^*(y_1(s), \dots, y_{k(s)}(s)) = \min_{\theta \in \Omega} \sum_{i=1}^{k(s)} y_i(s) t_i(k(s), \theta). \quad (1)$$

Среднее время одной пользовательской сессии в иерархическом меню  $H$  складывается из средних времен нахождения пользователя в каждой панели меню  $s$  иерархии  $H$ , умноженных на вероятности  $\mu_s$  попадания в эту панель меню:

$$T(H) = \sum_{s \text{ из } H} \mu_s t^*(y_1(s), \dots, y_{k(s)}(s)) \quad (2)$$

и задача состоит в поиске иерархии, минимизирующей время сессии (2) на множестве всех возможных древовидных иерархий, надстроенных над множеством  $N$ .

### 3. ЭЛЕМЕНТЫ ТЕОРИИ ОПТИМИЗАЦИИ ИЕРАРХИЧЕСКИХ СТРУКТУР

Задача минимизации среднего времени поиска в иерархическом меню принадлежит широкому классу задач поиска оптимальных иерархических структур [22, 23]. Проблемы поиска оптимальных иерархий встречаются во многих областях, но, несмотря на различные интерпретации задач, имеют общую математическую базу, что позволяет применить общий подход к их решению. Основы такого подхода были заложены в работах [22, 23].

В частности, доказано [23], что для критерия оптимизации вида (2), представимого в виде суммы однородных функций, оптимально так называемое «однородное дерево», в котором каждая категория имеет одинаковое число подкатегорий с популярностями, распределенными в одной и той же пропорции  $y_1, \dots, y_k$  ( $y_1 + \dots + y_k = 1$ ). Пример однородного дерева ширины 2 с пропорцией (1/3, 2/3) приведен на рис. 3.

Среднее время поиска в иерархии  $H$  ширины  $k$  с пропорцией  $y_1, \dots, y_k$  можно записать в виде

$$T(H) = - \sum_{w \in N} \mu(w) \ln \mu(w) \frac{t^*(y_1, \dots, y_k)}{k} - \sum_{i=1}^k y_i \ln y_i. \quad (3)$$

Показано [23], что задача поиска оптимальной древовидной иерархии сводится к вычислению параметров дерева, минимизирующих выражение (3)

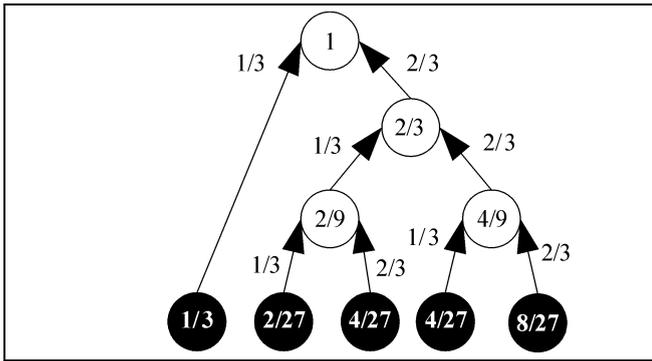


Рис. 3. Пример однородного дерева

по всем  $k = 2, \dots, n$  и пропорциям  $y_1, \dots, y_k$ , удовлетворяющим условию  $y_i \geq \min_{w \in N} \mu(w)$ :

$$T_L(N) = - \sum_{w \in N} \mu(w) \ln \mu(w) \min_k \min_{y_1, \dots, y_k} \frac{t^*(y_1, \dots, y_k)}{- \sum_{i=1}^k y_i \ln y_i}. \quad (4)$$

В силу дискретности параметров наилучшее однородное дерево зачастую невозможно построить точно, но среднее время поиска в нем (4) всегда может быть вычислено, и оно дает хорошую нижнюю оценку среднего времени поиска в оптимальной иерархии. Предложены [24] эффективные алгоритмы, позволяющие построить большое число субоптимальных деревьев на основе различных группировок элементов нижнего уровня<sup>5</sup>.

#### 4. ОПТИМАЛЬНЫЕ ИЕРАРХИЧЕСКИЕ МЕНЮ

##### 4.1. Общие выводы

Таким образом, в рамках рассматриваемой модели оптимальная структура меню должна представлять собой однородную иерархию, т. е. каждая панель меню должна содержать по возможности одно и то же число вариантов с популярностями, поделенными в одной и той же пропорции. Легко видеть, что вследствие этого применение нескольких типов меню в одной иерархии неэффективно. Действительно, согласно выражению (1), наилучший тип меню  $\theta \in \Omega$  для каждой панели  $s \subseteq N$  определяется минимизацией

$$\theta^*(s) = \arg \min_{\theta \in \Omega} \sum_{i=1}^{k_s} y_i(s) t_i(k(s), \theta). \quad (5)$$

<sup>5</sup> Это дает возможность строить с помощью данного алгоритма субоптимальные иерархические меню с учетом семантических ограничений — требований осмысленности каждой категории (см. далее § 5).

В силу однородности оптимальной иерархии, ширина меню  $k(s)$  и пропорция  $y_1(s), \dots, y_{k(s)}(s)$  одинаковы для всех панелей. Следовательно, минимум (5) будет достигаться в одной и той же точке для всех панелей. Другими словами, все панели оптимального меню одного и того же типа независимо от их положения в иерархии. Это дает возможность сначала найти параметры наилучшего однородного дерева для каждого из рассматриваемых типов меню  $\theta \in \Omega$ , а затем рекомендовать в качестве оптимального тип меню, однородное дерево для которого дает минимальное время сессии.

При построении удобного меню следует учесть множество факторов: вид панелей меню, стратегию поведения, квалификацию и личные качества пользователей. Даже в одном и том же меню различные пользователи могут вести себя по-разному. Так, варианты в меню можно просматривать последовательно или, при известной сортировке вариантов, пользоваться методом деления пополам. Пользователи, часто пользующиеся меню, могут помнить расположение необходимого варианта и сразу выбирать его (так называемое «экспертное» поведение). Кроме того, на скорость работы и число совершаемых ошибок влияют личные качества пользователей, например, возраст (см. [17]). Пользователи с отличающимся поведением описываются различными выражениями времени просмотра панели меню. При этом в силу линейности модели эти временные затраты могут быть усреднены по всей аудитории, и оптимальное в среднем для всей аудитории меню — это меню для пользователя с усредненными временными затратами. Далее везде подразумевается использование усредненных оценок параметров временных задержек. В следующих параграфах решается задача поиска оптимальной структуры меню для классических моделей поведения пользователей и рассматриваются практически важные усложнения модели.

##### 4.2. Исчерпывающий поиск

Стратегия исчерпывающего поиска (exhaustive search [21]) подразумевает, что пользователь сначала просматривает все варианты на панели меню и только после этого совершает выбор. Считается, что пользователь не совершает ошибок в выборе вариантов. В этом случае время выбора варианта на панели не зависит от расположения этого варианта и, соответственно, равно среднему времени нахождения пользователя в панели меню

$$t(k) = t_{resp} + t_{load}k + t_{read}k + t_{click}, \quad (6)$$

где  $t_{resp}$  — время ответа сервера,  $t_{load}$  — время загрузки одного пункта меню,  $t_{read}$  — время прочтения и восприятия одного пункта меню,  $t_{click}$  — время, требуемое на совершение выбора.

Оптимальные параметры однородного дерева (число вариантов  $k$  и пропорция  $y_1, \dots, y_k$ ) доставляют минимум функции (3), что с учетом выражения (6) эквивалентно минимизации функции

$$-\frac{(t_{resp} + t_{load}k + t_{click}) + t_{read}k}{\sum_{i=1}^k y_i \ln y_i}.$$

Легко проверить, что при любом фиксированном  $k$  оптимальна симметричная иерархия, т. е.  $y_i = 1/k, i = 1, \dots, k$ . Поэтому оптимальное число вариантов  $k$  в каждой панели меню находится в результате оптимизации по всем  $k = 2, 3, \dots$  функции

$$[(t_{resp} + t_{load}k + t_{click}) + t_{read}k] / \ln k. \quad (7)$$

**Пример 1.** Рассмотрим параметры, типичные для Интернет-каталогов:  $t_{resp} = 2$  с,  $t_{load} = 0,02$  с,  $t_{read} = 1$  с,  $t_{click} = 1$  с. Подставляя эти значения в функцию (7) и проводя численную оптимизацию, получим, что в оптимальном меню каждая панель должна содержать  $k = 5$  вариантов. При этом варианты на панели меню должны иметь равные популярности. ♦

Приведенные результаты согласуются с выражением для оптимальной ширины меню, полученной в классической статье [8]. Главное отличие полученного нами результата в том, что в работе [8] рассмотрение изначально ограничивается однородными симметричными деревьями. Выше же формально доказывается, что именно однородное симметричное дерево с шириной, определяемой из выражения (7), оптимально на множестве всевозможных (в том числе, неоднородных и асимметричных) древовидных иерархий. Кроме того, в статье [8] результаты были получены при ограничении одинаковой популярности всех элементов множества  $N$ , в настоящей же работе это ограничение не используется.

### 4.3. Последовательный поиск

Стратегия последовательного поиска (sequential search [21]) предполагает, что пользователь просматривает варианты на панели меню в определенной последовательности и совершает выбор как только встречает нужный вариант, не просматривая остальные варианты. При такой модели поведения время на выбор варианта в панели меню зависит от порядкового номера этого варианта. Среднее время, проведенное пользователем в панели меню, определяется усреднением по всем вариантам с учетом их относительных популярностей:

$$t(y_1, \dots, y_k) = t_{resp} + t_{load}k + t_{read} \sum_{i=1}^k iy_i + t_{click}. \quad (8)$$

Обозначим  $A(k) = (t_{resp} + t_{load}k + t_{click})/t_{read}$ , тогда среднее время, проведенное в панели меню,

$$t(y_1, \dots, y_k) = t_{read} \left( A(k) + \sum_{i=1}^k iy_i \right). \quad (9)$$

Время поиска в наилучшей однородной иерархии определяется минимизацией функции (3), что с учетом выражения (9) эквивалентно минимизации функции

$$A(k) + \sum_{i=1}^k iy_i - t_{read} \frac{\sum_{i=1}^k y_i \ln y_i}{k}$$

по всем  $k$  и пропорциям  $y_1, \dots, y_k$ .

При фиксированном  $k$  из условий первого порядка видим, что оптимальная пропорция удовлетворяет условию  $y_i = a(k)^{A(k)+i}, i = 1, \dots, k$ , где

$a(k)$  — решение уравнения  $\sum_{i=1}^k a^{A(k)+i} = 1$ . При под-

становке оптимальной пропорции в функции (10) получим выражение для оптимального числа вариантов  $k$  в панели меню:

$$k = \arg \min_{k'=2,3,\dots} 1/\ln(1/a(k')). \quad (11)$$

**Пример 2.** Для типичных параметров меню, приведенных в примере 1, подстановкой в формулу (11) получим, что при последовательной стратегии оптимальна более широкая иерархия с числом вариантов  $k = 13$ . Популярности вариантов в оптимальном однородном дереве распределены в пропорции  $y_1 \approx 0,27, y_2 \approx 0,20, y_3 \approx 0,15, \dots, y_{13} \approx 0,007$ , т. е. на первый из 13 просматриваемых пользователем вариантов должно приходиться больше четверти популярности, в то время как на последний — меньше 0,01. Среднее время (3) одной сессии пользователя для набора элементов  $N$  примерно равно  $3,25 \sum_{w \in N} \mu(w) \ln \mu(w)$ . ♦

Таким образом, при последовательном чтении пользователем вариантов в панели меню самые популярные элементы или категории следует размещать в начале, причем популярность вариантов убывает в геометрической прогрессии. В статье [8] результаты исчерпывающего поиска были обобщены и на последовательную стратегию, т. е. в качестве решения было предложено симметричное меню (с отличными от первой стратегии параметрами). Как показано в настоящем параграфе, это неоптимально. Так, для параметров, использованных в примерах 1 и 2, выигрыш оптимального асимметричного меню по сравнению со структурой, предложенной в статье [8], составляет более 11 %. Заметим, что асимметричные меню практи-



чески не исследуются в литературе, хотя существуют целые области (голосовые меню телефонных систем массового обслуживания, меню мобильных устройств), в которых последовательный перебор пунктов диктуется самой технологией реализации меню. В этих случаях анализ, проведенный в настоящем параграфе, особенно актуален.

Аналогично можно вычислить параметры оптимального меню и для более сложных стратегий поиска и типов меню (например, двухуровневых меню, см. рис. 2).

#### 4.4. Учет ошибок пользователя

Пользователь не всегда может правильно определить, какой категории принадлежит искомый элемент. Это приводит к ошибкам при навигации в меню. Если пользователь выбрал категорию по ошибке, то, как правило, он может подняться на уровень выше, в предыдущую панель меню, чтобы изменить свой выбор. Будем считать, что названия терминальных элементов меню сформулированы четко и пользователь может ошибиться только с выбором категории. Возможны два типа ошибок: пользователь может выбрать неверную подкатегорию в панели меню или пропустить искомую подкатегорию в корректной панели и выйти на уровень выше. Будем считать, что если пользователь ошибся в некоторой панели, то при повторном ее просмотре он будет аккуратнее и дважды не ошибется.

Поведение пользователя, совершающего ошибку первого типа, можно смоделировать следующим образом. Пользователь попадает в панель меню, просматривает варианты и совершает выбор (ошибочный) согласно своей модели поведения (последовательный или исчерпывающий поиск). В новой панели меню пользователь вынужден просмотреть все варианты, чтоб понять, что подходящего среди них нет. После этого он выходит обратно в первую панель меню и совершает корректный выбор согласно своей стратегии поведения, как если бы он не совершал ошибки.

При втором типе ошибки последовательность действий пользователя остается такой же, с той разницей, что пользователь затрачивает дополнительное время на полный просмотр и выход из корректного меню, в которое он в итоге вернется. Для простоты дальнейшего изложения будет вестись в применении к ошибкам первого типа.

Если пользователь совершает ошибку, он затрачивает дополнительное время, которое можно разделить на две части: время, которое тратится на ошибочный выбор в верхней панели меню (совершение ошибки), и время на просмотр всех вариантов в подкатегории и выход из нее (возврат к верному пути). Потери времени, связанные с совершением ошибок, определяются стратегией по-

ведения пользователя. Так, если пользователь ошибочно выбирает вариант  $j$  в панели меню с  $k$  вариантами, то при последовательном поиске потери времени на совершение ошибки

$$t_{mistake}^j = t_{resp} + t_{load}k + t_{read}j + t_{click}, \quad (12)$$

а при исчерпывающем поиске эти потери не зависят от выбираемого варианта:

$$t_{mistake}^j = t_{resp} + t_{load}k + t_{read}k + t_{click}. \quad (13)$$

Для возврата к верному пути пользователь вынужден просмотреть все пункты в подкатегории и сделать выбор «вернуться выше», затратив на это время

$$t_{recover} = t_{resp} + t_{load}k_{sub} + t_{read}k_{sub} + t_{click}, \quad (14)$$

где  $k_{sub}$  — число вариантов в ошибочной панели меню.

Пусть пользователь может совершить ошибку в панели меню с вероятностью  $p$ . При этом вероятность ошибочного выбора категории  $i$  равна  $py_i$ . Содержательно это означает, чем популярнее категория, тем выше вероятность выбрать по ошибке ее. Тогда вероятность того, что за время сессии пользователь ошибочно попадет в категорию  $s$ , равна  $p\mu_s$ . Таким образом, к среднему времени пребывания в панели добавляется время, связанное с тем, что пользователь может совершить ошибку, и время на то, что пользователь может ошибочно попасть в панель и будет вынужден просмотреть ее всю и выйти обратно. В сумме это определяет добавку к среднему времени нахождения пользователя в панели  $s$  в размере

$$\mu_s p t_{recover} + \mu_s \sum_{i=1}^k y_i p t_{mistake}^i.$$

Если пользователи придерживаются стратегии исчерпывающего поиска, то время (13) на совершение ошибки и время (14) на возврат корректному состоянию из текущей панели меню равны времени совершения любого выбора в этой панели меню, определяемому выражением (6). Таким образом, с учетом вероятности ошибки среднее время, проведенное пользователем в панели меню при исчерпывающем поиске,

$$t_s(k) = (1 + 2p)(t_{resp} + t_{load}k + t_{read}k + t_{click}). \quad (15)$$

При подстановке времени (15) в выражение для среднего времени (3) поиска в меню легко видеть, что время растет линейно при увеличении вероятности ошибки, но точка минимума не зависит от  $p$ . Другими словами, при исчерпывающем поиске вероятность ошибок пользователя не влияет на структуру оптимального меню.

В случае, когда пользователи придерживаются стратегии последовательного поиска, добавим ко

времени (8) дополнительное время (12) и (14) с учетом вероятности ошибки и после преобразования получим, что среднее время, проведенное пользователем в панели меню,

$$t_s(y_1, \dots, y_k) = (1 + 2p)(t_{resp} + t_{load}k + t_{click}) + pt_{read}k + (1 + p)t_{read} \sum_{i=1}^k iy_i \quad (16)$$

Аналогично, как и в п. 4.3, введем обозначение

$$B(k) = \frac{(1 + 2p)(t_{resp} + t_{load}k + t_{click}) + pt_{read}k}{(1 + p)t_{read}} = \left(1 + \frac{p}{1 + p}\right) A(k) + \frac{p}{1 + p} k$$

и перепишем время (16) в виде

$$t_s(y_1, \dots, y_k) = (1 + p)t_{read} \left( B(k) + \sum_{i=1}^k iy_i \right).$$

Тогда при минимизации (3) из условий первого порядка получим, что оптимальная пропорция удовлетворяет условию  $y_i = b(k)^{B(k) + i}$ ,  $i = 1, \dots, k$ ,

где  $b(k)$  — решение уравнения  $\sum_{i=1}^k b(k)^{B(k) + i} = 1$ .

Оптимальное число вариантов в меню определяется минимизацией по всем  $k' = 2, 3, \dots$  функции  $1/\ln(1/b(k'))$ .

**Пример 3.** Пусть пользователь совершает ошибку в панели меню с вероятностью  $p = 0,2$ . Тогда для параметров меню, использованных ранее в примерах 1 и 2, оптимально меню с числом вариантов  $k = 8$  в каждой панели меню. При этом пропорция для популярностей подкатегорий  $y_1 \approx 0,25$ ,  $y_2 \approx 0,19$ ,  $y_3 \approx 0,15$ , ...,  $y_8 \approx 0,05$ . Сравнивая с примером 2, видим, что учет вероятности ошибки пользователя приводит к применению более узких меню, позволяющих быстрее просмотреть ошибочное меню и вернуться к корректному состоянию. ♦

## 5. УЧЕТ СЕМАНТИЧЕСКИХ АСПЕКТОВ ПРИ ПОСТРОЕНИИ МЕНЮ

В описанной задаче оптимизации семантика не учитывается — формально допустимы любые группировки элементов, и каждая характеризуется только суммарной популярностью. На практике множество допустимых структур меню сильно сужается благодаря семантическим ограничениям. В первую очередь, это требование на осмысленное наполнение категорий в меню. Семантическое качество названий категорий оказывает существенное влияние на время выбора [25]. Однако решение задачи без семантических ограничений позволяет сделать общие выводы о структуре и виде оптимального меню, а также разработать инструменты для частичной или полной автоматизации процесса наполнения меню.

Нижняя граница временных затрат (4) остается актуальной и для задачи с семантическими ограничениями. Таким образом, в первую очередь можно сравнить ожидаемые временные затраты в текущей структуре меню с нижней границей (4). Если разница незначительна (не более 5%), то текущее меню можно считать приемлемым.

При построении новой структуры меню разработчик может воспользоваться оптимальным деревом как шаблоном, наполняя его осмысленным содержимым. Как уже упомянуто, приблизительно оптимальных иерархий достаточно много, и вероятность найти среди них иерархии с осмысленным разбиением элементов на категории достаточно велика.

На основе алгоритмов [24] для задачи с семантическими ограничениями легко построить алгоритм для нахождения субоптимального дерева, состоящего только из осмысленных категорий. Этот алгоритм предполагает существование классификации элементов по нескольким критериям или наличие некоторой функции для определения семантической близости элементов. Однако на сегодня в большинстве случаев построение меню не может быть полностью осуществлено автоматически, и для разработчика проверка осмысленности  $2^n - n - 1$  нетривиальных категорий элементов крайне трудоемка. Поэтому более реалистичным видится не формализовать семантические аспекты при построении меню, а положиться в этом на возможности и опыт разработчика меню.

**Пример 4.** Рассмотрим пример построения меню для реальной системы, иллюстрирующий описанные здесь идеи. Одна из платформ, где проблема оптимизации меню стоит особенно остро — мобильные телефоны. Так как экран мобильного телефона, как правило, сильно ограничен в размерах, иерархическое меню выступают практически единственной альтернативой организации доступа к множеству предоставляемых команд. Кроме того, размеры экрана, как правило, не позволяют охватить панель меню целиком, и пользователь вынужден просматривать варианты последовательно.

Рассмотрим меню отправки и приема сообщений мобильного телефона «Nokia 7510». На рис. 4, а приведена сокращенная структура оригинального меню. Популярности команд и категорий, приведенные на рисунке, основаны на анализе использования телефона и статистических исследованиях [3, 5]. Заметим, что структура меню сильно неоднородна: на верхнем уровне предлагается 14 вариантов, в то время как на втором уровне пользователь выбирает всего из 2—5 вариантов (последующие уровни опущены на иллюстрации). Значит, как показано в § 4, эту структуру возможно оптимизировать. Направление оптимизации определяется конкретными параметрами рассматриваемого меню.

Из экспериментов можно выяснить средние временные параметры: время переключения в новую панель меню не зависит от числа вариантов в меню и равно  $t_{resp} \approx 1$  с, среднее время чтения одного пункта  $t_{read} \approx 1$  с (включает в себя время перелистывания пун-

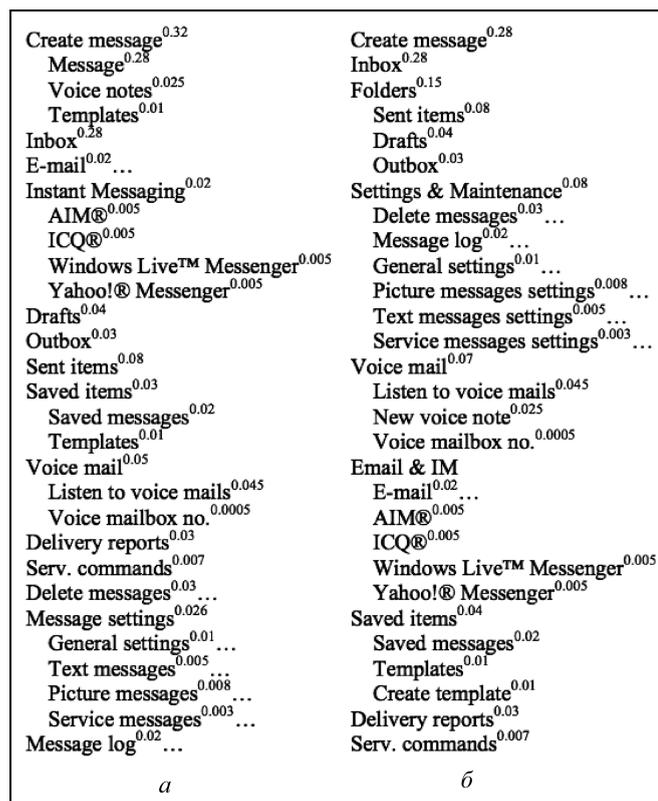


Рис. 4. Оригинальная (а) и оптимизированная (б) структуры меню мобильного телефона

ктов меню на экране), время нажатия на кнопку выбора  $t_{click} \approx 0,5$  с. Будем считать, что пользователи придерживаются стратегии последовательного поиска (неудобство средств управления очень ограничивает поведение пользователя) и с вероятностью  $p = 0,05$  могут совершить ошибку в панели меню. Тогда среднее время пребывания пользователя в панели меню выражается формулой (16). Для принятых параметров численный расчет показывает, что среднее время одной пользовательской сессии в оригинальном меню составляет  $T \approx 8,36$  с.

Согласно результатам п. 4.3 оптимальное меню содержит  $k = 8$  вариантов в каждой панели меню, а оптимальная пропорция для популярностей  $y \approx (0,33, 0,23, 0,16, 0,11, 0,07, 0,05, 0,03, 0,02)$ , т. е. структура оптимального меню сильно асимметрична. Теоретический минимум среднего времени доступа для рассматриваемого набора команд 6,74 с.

Однако построить оптимальное дерево невозможно в силу содержательных ограничений (команды меню нельзя организовать произвольным образом) и дискретности задачи. Используя алгоритмы построения субоптимальных деревьев [24] и учитывая содержательные ограничения, команды меню могут быть реорганизованы в структуру, приведенную на рис. 4, б. Численный расчет показывает, что среднее время пребывания пользователя в таком меню составляет примерно 7,22 с. Таким образом, изменение структуры меню согласно вычисленным оптимальным параметрам увеличивает скорость доступа к командам в среднем на 13,5 %. Срав-

нивая аналитический и численный результат, можно видеть, что построенное меню, хотя и не единственно возможное, близко по временным затратам к теоретическому минимуму. При этом построенное оптимизированное меню не нарушает логической структуры организации команд в категории. ♦

## 6. АВТОМАТИЗАЦИЯ ПОСТРОЕНИЯ МЕНЮ

На основе рассуждений, приведенных в § 5, можно предложить алгоритм для автоматизации процесса построения меню, опираясь на представленную теорию. Опишем основные шаги при проектировании структуры меню с применением САПР.

*Шаг 1.* Разработчик загружает в систему набор элементов  $N$  с их популярностями (полученными из экспериментов или основанных на экспертных оценках). Эти элементы в конечном итоге будут размещены в меню.

*Шаг 2.* Разработчик выбирает типы меню, применимые в рассматриваемом окружении, и уточняет время, требуемое на совершение элементарных действий, с учетом планируемого режима использования меню и квалификации потенциальных пользователей.

*Шаг 3.* Система рассчитывает нижние границы среднего времени пользовательской сессии для различных типов меню и предлагает наилучший тип.

*Шаг 4.* Разработчик выбирает тип меню, наиболее удобный и эффективный в рассматриваемом окружении, тем самым фиксируя оптимальное число пунктов  $k$  на панели и пропорцию  $y_1, \dots, y_k$  для их популярностей.

*Шаг 5.* Разработчик разбивает множество элементов  $N$  на осмысленные категории  $s_1, \dots, s_k$ , стремясь при этом приблизиться к оптимальному числу категорий и распределению популярностей для них. Система направляет разработчика, вычисляя качество текущего разбиения, например, по формуле

$$\frac{c(\mu(s_1), \dots, \mu(s_k)) - \mu(N)c(y_1, \dots, y_k)}{\mu(N)c(y_1, \dots, y_k)}. \quad (17)$$

*Шаг 6.* После того как получено приемлемое разбиение, процедура повторяется внутри каждой категории, и так далее до уровня, на котором останутся лишь отдельные элементы. Система при этом рассчитывает отклонение среднего времени пользовательской сессии от теоретически оптимального значения.

*Шаг 7.* При необходимости разработчик вручную улучшает полученную структуру, модифицируя категории с худшими оценками качества (17).

Процесс разбиения элементов на категории также может быть автоматизирован, если существует классификация элементов множества  $N$  (на-

пример, ключевые слова для сайтов или публикаций). Эту классификацию можно использовать для автоматического размещения элементов в иерархии с оптимальными (или близкими к оптимальным) параметрами.

### ЗАКЛЮЧЕНИЕ

Анализ предложенной модели показывает, что оптимальная структура иерархического меню представляет собой однородное дерево. Как правило, абсолютно однородное дерево не удастся построить. Но всегда существует множество иерархий, близких к оптимальным, и среди них есть меню с осмысленным разбиением элементов на категории. Представленные результаты позволяют получить оптимальные параметры меню, стремясь к которым можно построить эффективное меню, как показано в примере 4.

Описанная методика применима к широкому кругу представлений меню и стратегий поведения пользователей. Так, на основе базовых стратегий пользователей изучены варианты усложнения модели, в частности, учет ошибок пользователя. Общая модель позволяет доказать, что в одной иерархии выгоднее применять один тип для всех панелей меню.

Существует большое число экспериментальных исследований поведения пользователей в меню различных типов. Представленная общая модель фиксирует параметры, важные для среднего времени поиска, позволяя сфокусировать на них усилия в ходе экспериментов для дальнейшего прогнозирования и оптимизации структуры меню.

Полученные теоретические результаты могут быть применены в САПР для построения эффективного меню. Система на основе параметров внешней среды, выбранных пользователем, рассчитывает оптимальные свойства и тип меню и предоставляет удобные средства для заполнения меню по эффективному шаблону. При заполнении предоставляются средства для сравнения среднего времени доступа в полученном меню с теоретическим минимумом для оценки эффективности текущей структуры. Дальнейшая автоматизация предполагает автоматическое заполнение структуры меню на основе классификации элементов.

### ЛИТЕРАТУРА

1. Пархоменко П.П. Вопросники и организационные иерархии // Автоматика и телемеханика. — 2010. — № 6 (в печати).
2. Norman K.L. The Psychology of Menu Selection: Designing Cognitive Control at the Human/Computer Interface. — Norwood, N.J.: Ablex Publishing Corporation, 1991. — 368 p.
3. Thimbleby H. Analysis and Simulation of User Interfaces // Proc. BCS Human Computer Interaction. — 2000. — P. 221–237.
4. Thimbleby H. Press On: Principles of Interaction Programming. — The MIT Press, 2007. — 528 p.

5. Andersson E., Isaksson I.-M. Exploring alternatives to the hierarchical menu structure used in mobile phones. — Umea, 2007. — 70 p.
6. Landauer T.K., Nachbar D.W. Selection from alphabetic and numeric menu trees using a touch screen: depth, breadth and width // Proc. of the SIGCHI conf. on Human Factors in Computing Systems. — 1985. — P. 73–78.
7. Miller D.P. The depth/breadth tradeoff in hierarchical computer menus // Proc. of the Human Factors Society's 25th Annual Meeting. — 1981. — P. 296–300.
8. Lee E., MacGregor J. Minimizing User Search Time in Menu Retrieval Systems // Human Factors. — 1985. — Vol. 27, N 2. — P. 157–162.
9. Snowberry K., Parkinson S., Sisson N. Computer display menus // Ergonomics. — 1983. — Vol. 26. — P. 699–712.
10. Paap K.R., Roske-Hofstrand R.J. The Optimal Number of Menu Options per Panel // Human Factors. — 1986. — Vol. 28, N 4. — P. 377–385.
11. Cockburn A., Gutwin C., Greenberg S.A. Predictive Model of Menu Performance // Proc. ACM CHI'07. — 2007. — P. 627–636.
12. Sears A., Shneiderman B. Split menus: Effectively using selection frequency to organize menus // ACM ToCHI. — 2004. — Vol. 1, N 1. — P. 27–51.
13. Shneiderman B. Software Psychology: Human Factors in Computer and Information Systems. — Cambridge, Mass.: Winthrop Publishers, 1980. — 320 p.
14. Kiger J.I. The depth/breadth tradeoff in the design of menu-driven interfaces // International Journal of Man-Machine Studies. — 1984. — Vol. 20. — P. 201–213.
15. Jacko J.A., Salvendy G. Hierarchical menu design: breadth, depth, and task complexity // Perceptual and Motor Skills. — 1996. — Vol. 82. — P. 1187–1201.
16. Larson K., Czerwinski M. Web page design: Implications of memory, structure and scent from information retrieval // Proc. of the Association for Computing Machinery's CHI'98. — 1998. — P. 18–23.
17. Zaphiris P. Depth vs. breadth in the arrangement of web links // Proc. of the 44th Annual Meeting of the Human Factors and Ergonomics Society. — 2000. — P. 139–144.
18. Norman K.L., Chin J.P. The effect of tree structure on search in a hierarchical menu selection system // Behaviour and Information Technology. — 1988. — Vol. 7. — P. 51–65.
19. Bernard M.L. Examining a metric for predicting the accessibility of information within hypertext structures / Ph.D. thesis. — Wichita State University, 2002.
20. Roberts T.L., Engelbeck G. The Effects of Device Technology on the Usability of Advanced Telephone Functions // International Conference on Human Factors in Computing Systems CHI. — 1989. — P. 331–338.
21. Hollink V., Van Someren M., Wielinga B. Navigation behavior models for link structure optimization // User Modeling and User-Adapted Interaction. — 2007. — Vol. 17, N 4. — P. 339–377.
22. Воронин А.А., Мишин С.П. Оптимальные иерархические структуры. — М.: ИПУ РАН, 2003. — 214 с.
23. Губко М.В. Математические модели оптимизации иерархических структур. — М.: ЛЕНАНД, 2006. — 264 с.
24. Губко М.В. Алгоритмы построения субоптимальных организационных иерархий // Автоматика и телемеханика. — 2009. — №1. — С. 162–179.
25. Mehlenbacher B., Duffy T.M., Palmer J. Finding information on a menu: Linking menu organization to the user's goals // Journal of Human-Computer Interaction. — 1989. — Vol. 4. — P. 231–251.

Статья представлена к публикации членом редколлегии чл.-корр. РАН Д.А. Новиковым.

Губко Михаил Владимирович — канд. техн. наук, ст. науч. сотрудник, ☎ (495) 334-90-51, ✉ mgoubko@mail.ru,

Даниленко Александр Игоревич — вед. инженер, аспирант, ☎ (495) 334-90-51, ✉ danilenko.alexander@gmail.com,

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



# ВЗАИМОДЕЙСТВИЕ КЛИЕНТА И СЕРВЕРА В ИНТЕРНЕТ-СЛУЖБЕ ОБРАБОТКИ УДАЛЕННЫХ ПАКЕТОВ ФАЙЛОВ

Р.Э. Асратян

Рассмотрены принципы организации сетевого взаимодействия в Интернет-службе RFPS (службе удаленных пакетов наборов данных), предназначенной для поддержки распределенных систем управления. Акцент сделан на задаче восстановления взаимодействия после случайного или намеренного разрыва сетевого соединения без потери результатов обработки. Описано ее решение в службе RFPS, основанное на использовании удаленного пакета наборов данных в качестве основы для создания стабильного контекста взаимодействия, не разрушающегося в результате сетевых сбоев.

**Ключевые слова:** распределенная система управления, Интернет-служба, сетевое взаимодействие.

## ВВЕДЕНИЕ

Появление Интернета дало новый импульс разработкам территориально-распределенных систем управления, пользующихся Интернетом для организации взаимодействий между удаленными компонентами. Вполне естественно, что разработчики распределенных систем традиционно опираются на Интернет-технологии (сетевые службы и протоколы), составляющие основу Всемирной сети (HTTP, SOAP, RPC, RMI и т. п.) [1–3].

Общая проблема всех онлайн-Интернет-служб, основанных на семействе протоколов TCP/IP [2, 4], состоит в их неустойчивости к сетевым сбоям. Причина коренится в общей для них логике взаимодействия между клиентом и сервером, согласно которой исходные данные информационного запроса и результаты его обработки не связывает между собой ничто, кроме того, что они передаются по одному и тому же двустороннему TCP-каналу в рамках одного и того же сетевого соединения. Другими словами, сетевое соединение играет роль «контекста» взаимодействия — при его разрушении результаты обработки запроса оказываются потерянными для клиента.

По-видимому, при разработке этих служб и протоколов предполагалось, что клиентское приложение должно в этом случае или попросту повторить запрос заново или принять собственные меры, чтобы сохранить результаты запроса на сервере и впоследствии каким-то образом получить к

ним доступ. Понятно, что такой подход приводит к появлению множества «частных решений» этой проблемы, привязанных к особенностям и логике работы конкретных систем. В наиболее распространенных сетевых технологиях (таких, как технология Web-сервисов [5] или новая технология WCF [6]) акцент сделан скорее на достижении максимальной производительности и совершенствовании пользовательского интерфейса, чем на защите от сетевых сбоев.

Безусловно, распределенные системы управления предъявляют более высокие требования к устойчивости взаимодействия, чем, например, Web-форум или сетевая служба новостей, так как цена потери данных может оказаться весьма высокой. Эти требования еще более возрастают в системах, в которых один информационный запрос может обрабатываться не в одном сервере, а в цепочке серверов (т. е. в системах, опирающихся на архитектуру «клиент — сервер — сервер —...— сервер»). Если организация системы предполагает возможность вовлечения в обработку одного запроса сразу нескольких серверов, то проблема защиты от сетевых сбоев не может быть решена на уровне прикладных программ, так как последние, как правило, не участвуют в организации межсерверных взаимодействий. Другими словами, в этом случае данная проблема должна быть решена на уровне самой Интернет-службы, обеспечивающей взаимодействие между сетевыми узлами. В данной работе мы рассмотрим принципы организации

таких взаимодействий в Интернет-службе RFPS (Remote File Packets Service), предназначенной для сетевой поддержки распределенных систем управления [7, 8]. В § 1 приводятся краткие сведения об этой технологии (более подробную информацию о целях создания RFPS, принципах его организации, преимуществах и недостатках в сравнении с другими сетевыми технологиями можно найти в работе [9]).

### 1. КРАТКИЕ СВЕДЕНИЯ О RFPS

Служба RFPS представляет собой сетевую технологию, специально ориентированную на организацию информационных взаимодействий в распределенных системах и приложениях. Ее главные особенности:

- наличие средств межсерверного взаимодействия и межсерверной маршрутизации данных;
- возможность обработки клиентского запроса в цепочке серверов;
- устойчивость к сетевым сбоям.

Семантика RFPS основана на понятии удаленного пакета наборов данных — поименованного контейнера информации, способного перемещаться по сети от сервера к серверу и аккумулировать результаты обработки на разных сетевых узлах. Основные формы обработки включают в себя:

- создание (открытие) пакета на обслуживающем сервере;

- заполнение его данными;
- применение к нему одного или нескольких обработчиков (программ, активизируемых по вызову);
- передачу пакета в очереди на обработку к агентам (постоянно активным программам, работающим на отдельных узлах);
- временное перемещение пакета на обработку в другой сервер и продолжение обработки на другом сервере (может выполняться рекурсивно в цепочке серверов);
- считывание результатов обработки из пакета по сети;
- закрытие (уничтожение) пакета.

Перечисленные формы обработки можно комбинировать в разных сочетаниях. Создав удаленный пакет на каком-либо сервере, RFPS, клиент может последовательно (или параллельно) применить к нему несколько обработчиков и (или) последовательно поставить его в очередь на обработку к нескольким агентам, и (или) переместить его на другой RFPS-сервер для продолжения обработки. При этом пакет данных в RFPS создает устойчивый «контекст» взаимодействия, позволяющий продолжить обработку даже после случайного (или намеренного) разрыва сетевого соединения.

Все компоненты, с которыми имеет дело сервер RFPS (клиенты, агенты, обработчики и серверы), должны быть предварительно зарегистрированы в его конфигурационных данных с указанием необ-

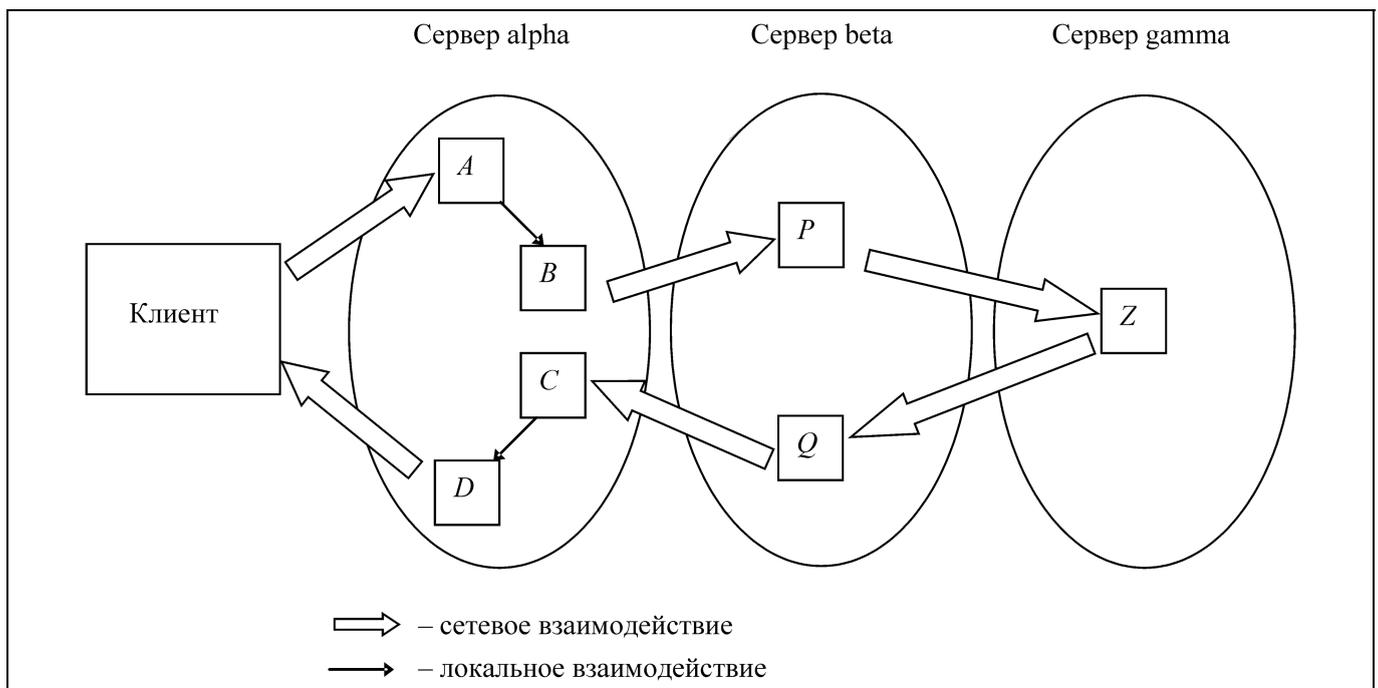


Рис. 1. Обработка пакета несколькими обработчиками на разных серверах



ходимых для работы характеристик (регистрационное имя и пароль клиента или агента, имя исполняемого модуля обработчика, Интернет-имя или IP-адрес сервера и т. п.). На рис. 1 проиллюстрирована обработка пакета в цепочке серверов с именами alpha, beta и gamma с помощью одного вызова клиентского метода «Process», имеющего два параметра: идентификатор пакета и командная строка, задающая последовательность вызовов обработчиков:

```
Client.Process(mypacket, «A, B, beta$(P, gamma$(Z, Q), C, D»).
```

Предполагается, что клиент соединен с сервером alpha, на котором создан обрабатываемый пакет (имя пакета сохранено в переменной mypacket). С помощью стрелок отображена последовательность вызова обработчиков, которые обозначены квадратиками. Выполнение вызова включает в себя:

- передачу пакета на обработку обработчикам *A* и *B* на сервере alpha;
- перемещение пакета на сервер beta и применение к нему обработчика *P*;
- перемещение пакета на сервер gamma, применение к нему обработчика *Z* и возврат пакета на сервер beta;
- применение к пакету обработчика *Q* на сервере beta и возврат пакета на сервер alpha;
- применение к пакету обработчиков *C* и *D* на сервере alpha.

Предполагается, что имена alpha, beta и gamma — регистрационные имена серверов, а *A*, *B*, *C*, ... — регистрационные имена обработчиков на соответствующих серверах.

## 2. ОРГАНИЗАЦИЯ ВЗАИМОДЕЙСТВИЯ МЕЖДУ КЛИЕНТОМ И СЕРВЕРОМ

Рассмотрим следующий фрагмент кода на языке C++, иллюстрирующий обращение к удаленному серверу со стороны клиента.

```
RFPSClient Rfppc; //Объект класса RFPSClient
char PackId[9]; //Строка для хранения
//идентификатора пакета
Rfppc.Config("192.168.1.77", 8126, "alibaba", "sezam");
//Параметры соединения
Rfppc.Connect(); //Соединение с RFPS-сервером
Rfppc.Open(PackId); //Открытие пакета
Занесение файлов в пакет
Rfppc.SendFile(PackId, "c:\\out", "file1.zip");
Rfppc.Process(PackId, "unzip,beta$trans(\\win-koil"),
zip", 300); //Обработка пакета
while(Rfppc.GetFile(PackId, "c:\\in", "*"*)==0);
//Считывание файлов из пакета
Rfppc.Close(PackId); //Закрытие пакета
Rfppc.Disconnect(); //Закрытие соединения
```

Первые две строки определяют две переменные: строку символов PackId и объект класса RFPSClient. Этот класс содержит в себе все методы, необходимые клиенту для работы с RFPS. Третья строка настраивает объект на работу с определенным RFPS-сервером. В качестве параметров объекту передаются IP-адрес (или имя) сервера, номер порта, имя и пароль пользователя (RFPS-сервер обслуживает только зарегистрированных пользователей).

Четвертая строка устанавливает TCP-соединение с сервером (на сервере немедленно создается отдельный процесс для обслуживания этого соединения), а пятая — открывает на сервере новый пакет: в результате выполнения метода Open строка PackId окажется заполненной уникальным (для конкретного сервера) идентификатором открытого пакета. Все последующие методы будут пользоваться этим идентификатором.

Шестая строка обеспечивает передачу на сервер файла из каталога «c:\\out». Можно представлять себе, что файл помещается по сети в открытый пакет. Седьмая строка иницирует обработку путем вызова метода Process. В качестве аргумента передается командная строка, обеспечивающая последовательный вызов трех обработчиков. Причем первый и последний (зарегистрированные под именами «unzip» и «zip») выполняются непосредственно на обслуживающем сервере, а второй («trans») — на сервере с именем «beta». Обработчику «trans» передается строковый параметр «win-koil», который становится доступным обработчику как параметр командной строки. Наборы данных пакета становятся доступны обработчику как файлы в текущем каталоге. Последний параметр метода Process — размер таймаута (в секундах), ограничивающий предельное время обработки.

Восьмая строка содержит цикл выборки всех файлов из пакета в локальный каталог клиента. Девятая и десятая обеспечивают закрытие пакета и прекращение соединения с сервером соответственно.

Разумеется, из данного фрагмента кода намеренно удалены операторы проверки успеха выполнения методов и обработки исключений. Важно подчеркнуть, что идентификатор открытого пакета является обязательным параметром всех последующих методов, связанных с его обработкой.

Взаимодействие клиента и сервера на уровне TCP-соединения организуется по следующим правилам.

Сервер любой службы, основанной на семействе протоколов TCP/IP, постоянно занят «прослушиванием» входящих TCP-соединений на выделенном для него порте [4]. Зафиксировав входящее

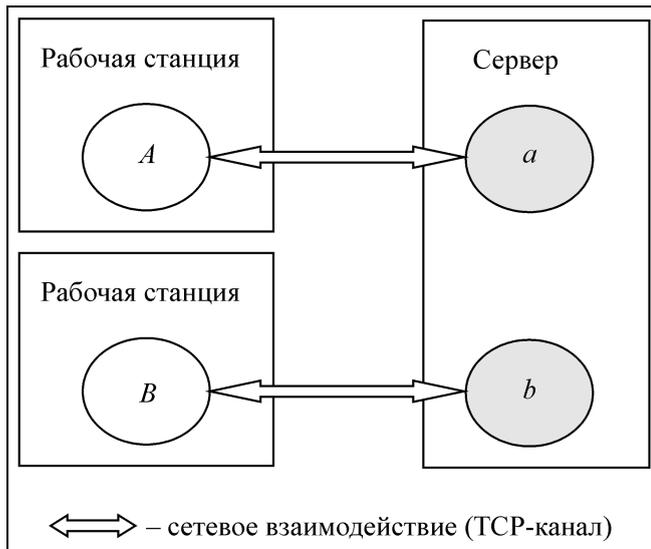


Рис. 2. Клиентские и обслуживающие процессы

соединение, сервер порождает отдельный обслуживающий процесс для обработки входящего запроса, создает двунаправленный TCP-канал между клиентским процессом (далее просто клиентом) и обслуживающим процессом и немедленно возвращается к прослушиванию входящих соединений. (Говоря более строго, для обработки запроса обычно порождается параллельная программная нить — `program thread`, разделяющая адресное пространство сервера, но в дальнейшем мы будем для краткости пользоваться термином «процесс».) Вся содержательная обработка запроса (получение кода операции, параметров и входных данных от клиента по TCP-каналу, собственно обработка запроса и возврат результатов клиенту) производится в обслуживающем процессе. (Например, в случае RFPS именно обслуживающие процессы обеспечивают создание и уничтожение пакетов наборов данных, вызов обработчиков, организацию обращения к удаленным серверам и т. п.) Данные правила проиллюстрированы на рис. 2, где клиентские процессы отображены белыми кружками и обозначены прописными буквами, а обслуживающие — темными кружками и строчными буквами.

Передав код операции и параметры обращения (и, возможно, данные для обработки) в созданный TCP-канал, клиент приступает к чтению результатов обработки из того же TCP-канала. Если результаты еще не готовы, клиент вновь переводится в неактивное состояние до появления результатов в TCP-канале. Если в процессе обработки произойдет неожиданное разрушение TCP-канала, операция чтения данных из канала вернет аномальный код завершения. Обслуживающий процесс на сер-

вере будет продолжать свою работу до того момента, когда он попытается записать в канал результаты обработки. В результате он также получит аномальный код завершения операции записи в канал, и ему не останется ничего другого, как просто завершить свою работу. Другими словами, в этом случае результаты обработки могут оказаться потерянными, если не принять специальных мер по их сохранению и обеспечению доступа к ним.

Рассмотрим как решается эта задача в RFPS на примере выполнения метода `Process`. В описанную схему взаимодействия клиента и сервера добавляется новая функциональность.

- При создании нового пакета в базе данных сервера RFPS создается новая запись — дескриптор пакета. В этой записи поддерживается идентификатор и управляющая информация пакета: имя создавшего клиента и дата создания, текущее состояние (обрабатывается или свободен), код завершения и системное диагностическое сообщение последней операции обработки, трасса (история) перемещения и обработки пакета, имена файлов, входящих в пакет и т. п. Эта запись удаляется только при закрытии пакета клиентом (или процедурой «сборки мусора», автоматически уничтожающей все слишком давно не обрабатываемые пакеты). Собственно наборы данных пакета размещаются в файловой системе сервера.

- При выполнении операции `Process` в памяти сервера RFPS создается дескриптор сеанса обработки, который содержит идентификатор обрабатывающего процесса в серверной ОС, время начала обработки, идентификатор пакета и несколько системных компонентов для синхронизации процесса обработки с другими процессами сервера (в реализации для платформы Win32 используются дескрипторы событий, а в реализации для UNIX — условные переменные). Сервер RFPS ведет учет всех открытых пакетов, сеансов и процессов таким образом, что каждый дескриптор сеанса и обслуживающий процесс ассоциируется с переданным ему (через параметры запроса) идентификатором пакета. В данном случае эта ассоциация означает, что по идентификатору пакета можно быстро определить адрес дескриптора сеанса обработки. (Разумеется, все это реализуется с помощью довольно сложной структуры данных сервера RFPS, включающей в себя связанные между собой дескрипторы сеансов, обслуживающих процессов и пакетов, но мы не будем останавливаться на описании этой структуры.)

- После завершения обработки каждого запроса обслуживающий процесс помещает код завершения обработки и системное диагностическое сообщение в дескриптор обрабатываемого пакета



и, кроме того, пытается передать их клиенту по TCP-каналу. Независимо от успеха этой передачи (соединение с клиентом может оказаться разорванным) обслуживающий процесс посылает сигнал активизации всем процессам, которые ждут окончания обработки пакета (если они имеются) и немедленно заканчивает работу.

- Если при издании запроса процесс-клиент обнаруживает, что соединение с сервером разрушено, он автоматически восстанавливает его (с рождением нового процесса на сервере). Если разрыв соединения произошел в период ожидания кода завершения от обслуживающего процесса, процесс-клиент немедленно издает новый специальный запрос к серверу RFPS: «Ожидание окончания обработки» (с помощью специального клиентского метода Wait) и возвращается в состояние ожидания кода завершения. (Это поведение заложено в клиентские методы класса RFPSClient и реализуется автоматически).

- Запрос «Ожидание окончания обработки» порождает новый процесс на сервере RFPS — процесс ожидания — и новый TCP-канал. Этот процесс прежде всего проверяет состояние пакета по его дескриптору. Если обработка пакета уже завершена, то процесс ожидания попросту считывает его код завершения и диагностическое сообщение из дескриптора сеанса пакета, передает этот код процессу-клиенту по TCP-каналу и заканчивает свою работу. В противном случае он определяет адрес дескриптора сеанса обработки и переходит в состояние ожидания завершения обработки пакета. Эта операция может итеративно повторяться несколько раз: если процесс ожидания также потеряет связь с процессом клиентом, то будет по-

рожден новый процесс ожидания и т. д. На рис. 3 изображена ситуация, когда в рамках сеанса обработки порождено несколько процессов (обозначены серыми кружками), один из которых (с) — обслуживающий, а остальные (w) — процессы ожидания.

- После завершения обработки обслуживающий процесс проверяет, имеются ли процессы ожидания, связанные с тем же сеансом связи. Если есть, он посылает сигнал активизации всем этим процессам и заканчивает свою работу. Тот из процессов ожидания, который имеет связь с процессом-клиентом, передает ему код завершения обслуживающего процесса по TCP-каналу, после чего все процессы ожидания, порожденные в рамках данного сеанса обработки, также заканчивают свою работу.

### 3. ОРГАНИЗАЦИЯ ВЗАИМОДЕЙСТВИЯ МЕЖДУ АГЕНТОМ И СЕРВЕРОМ

Агенты, как и обработчики, предназначены для обработки пакетов. Главное отличие агентов от обработчиков заключается в следующем:

- агенты представляют собой постоянно активные или периодически запускаемые программы (а не запускаются «по вызову»);

- агенты могут работать на отдельной рабочей станции и взаимодействовать с RFPS-сервером по сети (в этом отношении агенты подобны клиентам);

- агенты получают данные для обработки из входной очереди пакетов.

Отправка пакета на обработку агенту осуществляется с помощью метода Process: если в командной строке имеется имя агента, то обрабатываемый пакет переносится во входную очередь этого агента, а выполнение метода Process задерживается до возврата пакета из очереди.

Как и программа клиента, программа агента должна связываться с обслуживающим сервером по сети (фактически, агентов можно рассматривать как клиентов, наделенных дополнительной функциональностью). Основу программы агента должен составлять цикл последовательной выборки пакетов из входной очереди (с помощью специальных методов GetCount и Select) и их обработки. Обработка каждого отдельного пакета включает в себя сетевой обмен данными между рабочей станцией агента и RFPS-сервером (например, с помощью методов GetFile и SendFile). Подчеркнем, что метод Select обеспечивает получение лишь идентификатора пакета, но не данных. В ка-

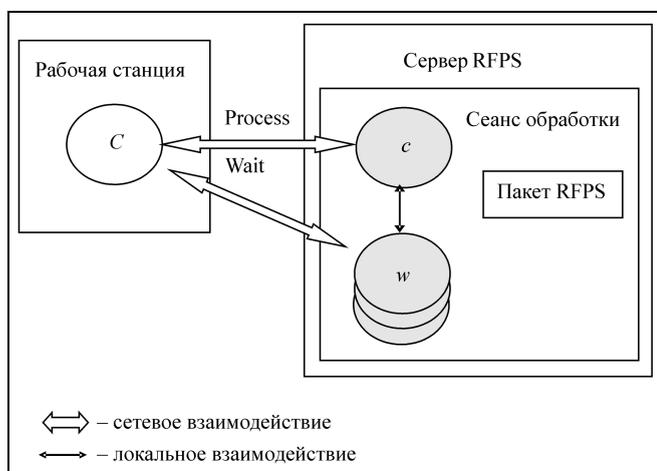


Рис. 3. Обслуживающий процесс и процессы ожидания

честве примера рассмотрим «скелет» исходного кода программы агента на языке C++:

```

RFPSClient Rfppc; //Объект класса RFPSClient
int QueLen; //Переменная для числа пакетов
//в очереди
char PackId[9]; //Строка для хранения
//идентификатора пакета
Rfppc.Config("192.168.1.77", 8126,"agent_007",
"sezam"); //Параметры соединения
Rfppc.Connect(); //Соединение с RFPS-сервером
while(1) //Начало «вечного» цикла обслуживания
//очереди
{
    Rfppc.GetCount(QueLen, 180); //Ожидание
    //пакетов в очереди (до 3 мин)
    if(QueLen<=0) continue; //Если их нет — снова
    //ожидание
    if(Rfppc.Select(1, PackId)==0) //Получение ид.
    //первого пакета в очереди
    {
        while(Rfppc.GetFile(PackId, "c:\\inp", "*.*")==0);
        //Считывание файлов из пакета
        ... //Обработка
        Rfppc.Return(PackId); //Возвращение пакета
        //клиенту-отправителю
    }
}
    
```

Как видно из примера, основу программы составляет вечный цикл (предложение while), в котором осуществляется:

- ожидание появления пакетов в очереди;
- последовательная выборка всех пакетов из очереди в рабочую область агента (с получением идентификатора каждого пакета);
- считывание файлов из каждого пакета;
- обработка данных пакета (ее исходный текст не приводится в примере);
- возврат пакета клиенту-отправителю (с автоматической активизацией его обслуживающего процесса).

Существует некоторая аналогия между очередью пакетов и почтовым ящиком электронной почты. В данном примере отражено их важное различие, связанное с возможностью синхронизации программы агента с событием появления пакетов во входной очереди с помощью метода GetCount. Данный метод возвращает число пакетов в очереди (в первом параметре), если она не пуста, или переводит программу агента в состояние ожидания в противном случае. Ожидающая программа будет автоматически активизирована после появления пакетов в очереди или после истечения таймаута (второй параметр).

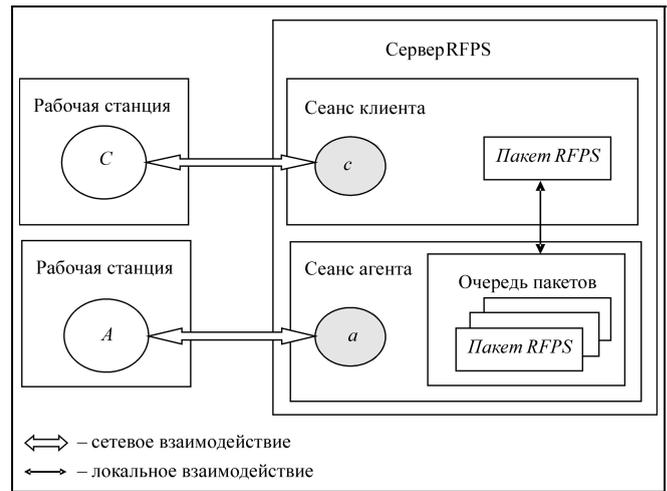


Рис. 4. Обработка пакетов агентом

Важно подчеркнуть, что агент — активное звено, а сервер — пассивное. Агент периодически обращается к обслуживающему серверу для выборки пакетов из собственной (поддерживаемой на сервере) входной очереди. Если агент потерял работоспособность, то входная очередь перестанет обслуживаться. По истечении заданного клиентом таймаута (размер таймаута является опциональным параметром метода Process) обслуживающий клиента процесс будет уничтожен на сервере, а клиент, обратившийся к агенту, получит сообщение о таймауте операции.

Как и в метод Process, методы GetCount и Select опираются на устойчивый контекст сеанса связи, но в данном случае он связывается не с пакетом, а с очередью пакетов. При обработке этих методов нет необходимости в «процессах ожидания» — в случае разрыва сетевого соединения обращение к серверу попросту повторяется. На рис. 4 проиллюстрирована передача пакета RFPS в очередь агента и обратно. Белыми кружками обозначены процессы клиента и агента (C и A соответственно), а темными — обслуживающие процессы на сервере RFPS.

Наконец, несколько слов о выполнении методов SendFile и GetFile, предназначенных для обменов данными между пакетом RFPS и файловой системой клиента или агента (и их аналогов SendMem и GetMem, предназначенных для обменов между пакетом и оперативной памятью). В случае разрыва сетевого соединения при передаче набора данных по сети данные методы обеспечивают автоматическое восстановление соединения и продолжение обмена с «точки разрыва», а не с начала набора данных. (При этом применяется примерно та же технология восстановления обмена,



что и в протоколе FTP [1, 2], которая основана на передаче дополнительного параметра — смещения в наборе данных — от узла-получателя в узел-отправитель.)

### ЗАКЛЮЧЕНИЕ

Семантика RFPS построена таким образом, чтобы все основные сетевые операции, связанные с передачей и удаленной обработкой данных, выполнялись в рамках стабильного контекста, независимого от сетевого соединения. В зависимости от операции, этот контекст связывается или с удаленным пакетом наборов данных, или с очередью пакетов. Это свойство RFPS избавляет прикладного программиста от необходимости разрабатывать собственные средства защиты от сетевых сбоев, что особенно важно в аспекте организации сложных межсерверных взаимодействий.

В статье не рассмотрена организация взаимодействия между серверами при обработке пакета в цепочке серверов. Отметим лишь, что оно построено примерно по тем же правилам, что и взаимодействие «клиент — сервер», так как при обращении к удаленному серверу вызывающий сервер выступает «в роли» клиента. И в этом случае контекст взаимодействия связывается с пакетом наборов данных, а не с сетевым соединением между серверами.

К недостаткам RFPS следует в первую очередь отнести относительно низкий уровень пользовательского интерфейса и отсутствие формализованных спецификаций обработчиков и агентов. В этом отношении он, безусловно, проигрывает, например, технологии Web-сервисов [5], опирающейся на формализованное описание удаленных процедур на языке WSDL и поддержанной в ряде распространенных систем программирования.

Одно из эффективных направлений применения RFPS состоит в создании «прозрачных» RFPS-туннелей между HTTP-клиентами и HTTP-серверами [10]. Суть подхода заключается во временной инкапсуляции HTTP-запроса и HTTP-ответа в пакет RFPS при прохождении данных через глобальную сеть в ситуации, когда HTTP-клиент и HTTP-сервер расположены в удаленных друг от друга частных сетях. Этот подход позволяет в значительной степени совместить преимущества технологии Web-

сервисов (высокоуровневый пользовательский интерфейс) и преимущества RFPS (устойчивость к сетевым сбоям и наличие встроенных средств межсерверного взаимодействия и межсерверной маршрутизации данных). В процессе лабораторных экспериментов с RFPS, в частности, было проведено встраивание RFPS-туннелей в несколько ранее разработанных распределенных систем, построенных на основе технологии Web-сервисов. (Подход может быть применен и для систем, построенных на новой технологии WCF [6], при условии, что в качестве «транспорта» выбран HTTP.) Важно отметить, что ни в одном случае не понадобилась какая-либо программная адаптация этих систем, что подтверждает «прозрачность» туннелей для HTTP-клиентов и серверов.

### ЛИТЕРАТУРА

1. Фролов А.В., Фролов Г.В. Глобальные сети компьютеров. — М.: Диалог-МИФИ, 1996. — 256 с.
2. Шиндер Д.Л. Основы компьютерных сетей. — СПб.: Вильямс, 2003. — 656 с.
3. Гофф М. Сетевые распределенные вычисления: достижения и проблемы. — М.: КУДИЦ-ОБРАЗ, 2005. — 320 с.
4. Джамса К., Коуп К. Программирование для Интернет в среде Windows. — СПб.: «Питер», 1996. — 659 с.
5. Шапошников И.В. Web-сервисы Microsoft .NET. — СПб.: БХВ-Петербург, 2002. — 336 с.
6. Лева Д. Создание служб WCF. — СПб: Питер, 2008. — 592 с.
7. Асратян Р.Э. Интернет-служба для поддержки распределенных информационно-управляющих систем // Проблемы управления. — 2005. — № 6. — С. 73—77.
8. Асратян Р.Э. Интернет-служба для поддержки межсерверных взаимодействий в распределенных информационных системах // Проблемы управления. — 2006. — № 5. — С. 58—62.
9. Асратян Р.Э., Лебедев В.Н. Средства информационного взаимодействия в современных распределенных гетерогенных системах. — М.: ЛЕНАНД, 2008. — 120 с.
10. Асратян Р.Э. Межсерверная маршрутизация HTTP/SOAP-взаимодействий в распределенных системах // Проблемы управления. — 2008. — № 5. — С. 57—61.

Статья представлена к публикации руководителем РРС В.Ю. Столбовым.

Асратян Рубен Эзрасович — канд. техн. наук, вед. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ☎ (495) 334-88-61, ✉ rea@ipu.ru.

# ОПТИМИЗАЦИЯ РАЗВИТИЯ СИСТЕМООБРАЗУЮЩЕЙ ЭЛЕКТРИЧЕСКОЙ СЕТИ С ПОМОЩЬЮ ГЕОИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ<sup>1</sup>

О.М. Попова, И.Ю. Усов

Дана методика решения задачи оптимизации развития системообразующей электрической сети электроэнергетической системы. Представлен состав соответствующего программно-вычислительного комплекса. Приведены показатели структурного анализа электроэнергетических систем. Показаны особенности применения геоинформационных технологий.

**Ключевые слова:** электроэнергетическая система, оптимизация развития, структурный анализ, геоинформационная система.

## ВВЕДЕНИЕ

Исторически развитие электроэнергетики связано с формированием электроэнергетических систем (ЭЭС), представляющих собой совокупность объектов (генераторов, трансформаторов, линий электропередач, установок электропотребителей, средств регулирования и управления), которые предназначены для производства, передачи, распределения и использования электрической энергии. Оно в значительной степени зависит от сооружения системообразующей электрической сети (СЭС), которая выполняет три главные функции: выдачу мощности крупных электростанций, электроснабжение крупных узлов нагрузки (промышленных агломераций, мегаполисов и др.) и осуществление совместной работы энергосистем в составе Единой энергетической системы России. Элементами СЭС служат, как правило, линии электропередачи и трансформаторы напряжением 330 кВ и выше.

При обосновании развития СЭС электроэнергетических систем формируется совокупность наиболее экономичных вариантов ее развития, осуществляется их технико-экономическое сопоставление и выбирается наилучшее решение для последующе-

го рабочего проектирования, строительства и ввода в эксплуатацию сетевых объектов.

С учетом территориально-временного аспекта управления развитием ЭЭС [1, 2] выделяются две основные последовательно решаемые задачи:

- оптимизация перспективной структуры СЭС с формированием наиболее экономичных вариантов ее развития и общей оценкой ресурсов (финансовых, технических и др.), требуемых для их реализации;
- технико-экономическое сравнение и выбор наиболее рациональных схем развития СЭС.

Первая из этих задач решается на уровне Единой национальной электрической сети (ЕНЭС) и СЭС объединенных энергосистем при сформированной структуре генерирующих мощностей на перспективу 15–20 лет. Вторая задача решается на межрегиональном и региональном уровнях на основе уточненной исходной информации с привлечением оценочных экономических моделей развития и детальных моделей физических процессов ЭЭС (расчет режимов, устойчивости, надежности) на перспективу 7–15 лет.

Поскольку методическое обеспечение решения второй задачи достаточно разработано [3–5], представленный в настоящей работе методический подход ориентирован на решение первой задачи.

В разные годы у нас в стране и за рубежом был разработан ряд математических моделей оптимизации развития электрической сети [6–11]. Эти

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (грант № 10-07-00264).



модели учитывают дискретный характер ввода новых элементов сети, и поэтому имеют ограничение на число рассматриваемых в них новых связей (линий электропередач). Еще один недостаток заключается в упрощенном учете ограничений на пропускную способность ветвей графа, посредством которого моделируется развивающаяся электрическая сеть. В действительности пропускная способность конкретной линии электропередачи определяется не только ее конструктивными характеристиками, но и условиями статической устойчивости [12]. Эти условия для каждой линии зависят от параметров сети в целом, распределения мощности между электростанциями и других факторов.

Представленный в работе методический подход позволяет рассмотреть эффективность сооружения практически неограниченного числа новых элементов электрической сети, а также более точно, чем в существующих моделях, учесть ограничения на передаваемые в сети потоки мощности по условиям статической устойчивости.

Этот подход предполагает линейную потоковую модель оптимизации развития электрической сети в качестве основного инструмента формирования наиболее рациональных вариантов развития СЭС [11]. Применение такой модели для решения практических задач оптимизации СЭС на длительную перспективу вполне обосновано, поскольку фактор дискретности в данном случае не выступает определяющим.

Для учета условий перспективного функционирования ЭЭС (ограничений на передаваемые в электрической сети мощности по условиям статической устойчивости) применяются алгоритмы структурного анализа ЭЭС [13, 14], реализованные в виде компьютерной программы.

Кроме того, на единой платформе с линейной потоковой моделью разработан геоинформационный блок программно-вычислительного комплекса [11, 15]. Для эффективного решения задач развития ЭЭС необходимо средство, позволяющее естественным и наглядным образом обрабатывать и анализировать разнородную и пространственно-координированную информацию. Таким средством на настоящий момент времени служат геоинформационные системы (ГИС), которые как специализированные системы обработки пространственно-временных данных, а также геоинформационные технологии изучаются новым активно развивающимся направлением информатики — геоинформатикой [16, 17]. Особенности построения ГИС развивающихся электрических сетей как проблемно-ориентированной системы рассмотрены в работе [15]. В настоящей статье сделан акцент на использовании в составе программно-вычислительного комплекса геоинформационных технологий, понимаемых прежде всего как технологий сбора, хранения, обработки и представления данных [18, 19]. С позиций управления геоинформационные технологии предполагают проектирование и поддержку принятия решений. Поскольку используемая информация в основном представляется в графической форме, то эти технологии близки к технологиям визуальной обработки информации.

Визуализация существенно облегчает восприятие проектировщиком схем сети, обеспечивает учет территориальных и структурных особенностей СЭС при определении принципов ее развития, а также позволяет осуществлять контроль правильности формирования схем, определение протяженности электрических связей и поиск соответствующих трасс прокладки линий электропередач.

Визуализация существенно облегчает восприятие проектировщиком схем сети, обеспечивает учет территориальных и структурных особенностей СЭС при определении принципов ее развития, а также позволяет осуществлять контроль правильности формирования схем, определение протяженности электрических связей и поиск соответствующих трасс прокладки линий электропередач.

## 1. ПОКАЗАТЕЛИ СТРУКТУРНОГО АНАЛИЗА ЭЛЕКТРОЭНЕРГЕТИЧЕСКОЙ СИСТЕМЫ

Поскольку рассмотрение детальных моделей электрических режимов в процессе оптимизации развития СЭС затруднено из-за большого числа вариантных расчетов, то для учета ограничений на передаваемые в электрической сети мощности по условиям статической устойчивости можно воспользоваться следующими показателями структурного анализа электроэнергетических систем:

— взаимные структурные мощности генераторов [13]:

$$W_{ij} = E_i E_j y_{ij}, \quad (1)$$

где  $E_i, E_j$  — переходные ЭДС генераторов в модели «шины — переходное сопротивление — переходная ЭДС» [12], а  $y_{ij}$  — взаимная проводимость между узлами, характеризуемых ЭДС  $E_i$  и  $E_j$ ;

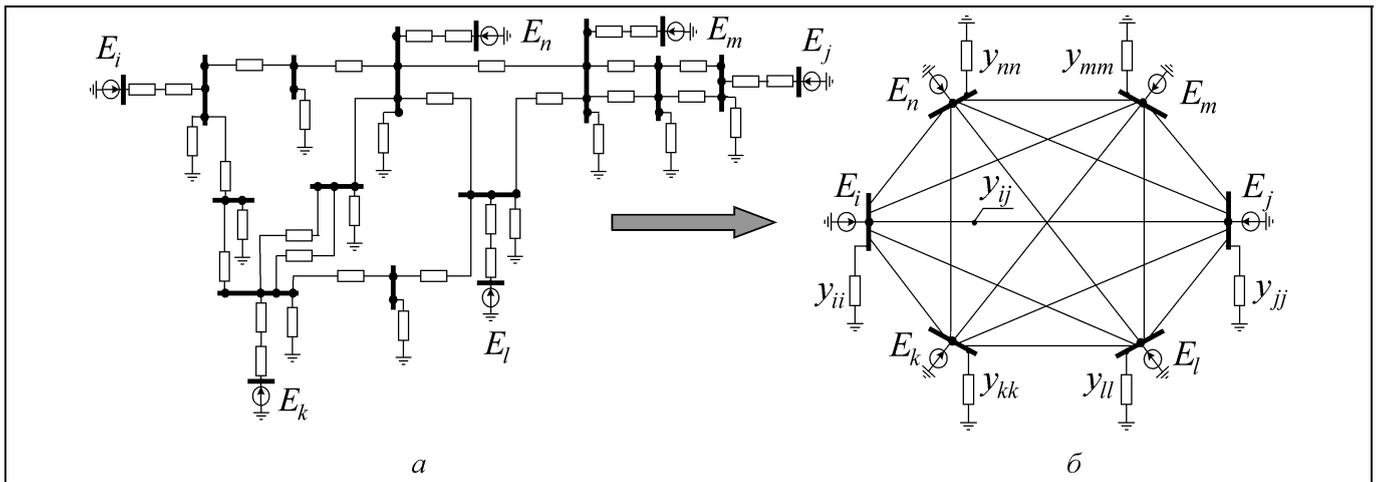
— собственные структурные мощности генераторов [13]:

$$W_{ii} = E_i^2 g_{ii}, \quad (2)$$

где  $g_{ii}$  — активная составляющая собственной проводимости узла  $i$  ( $y_{ii}$ ).

Соотношения токов и напряжений схемы замещения электрической сети (рис. 1, а), в которой ее элементы представлены своими электрическими параметрами (проводимостями и ЭДС), описываются системой уравнений узловых напряжений [12]

$$[Y] \begin{bmatrix} \dot{E} \\ \dot{U} \end{bmatrix} = \begin{bmatrix} \dot{I}_r \\ 0 \end{bmatrix}, \quad (3)$$


**Рис. 1. Схема замещения ЭЭС:**

*a* — расчетная; *б* — эквивалентная (взаимные проводимости узлов с ЭДС изображены сплошными линиями)

где  $[\dot{Y}]$  — матрица собственных и взаимных проводимостей (СВП) узлов,  $\dot{U}$  — вектор напряжений в узлах,  $\dot{E}$  — вектор переходных ЭДС в генераторных узлах,  $\dot{I}_r$  — вектор-столбец генерируемых токов в узлах.

Для определения показателей (1) и (2) схема замещения электрической сети, описываемая системой уравнений (3), приводится любым из известных методов эквивалентирования к полному графу, в вершинах которого лежат узлы с ЭДС (рис. 1, б). Эквивалентирование (преобразование исходной схемы в эквивалентную в целях упрощения дальнейшей работы с ней) осуществляется в данном случае для определения эквивалентной матрицы СВП относительно узлов с ЭДС —  $[\dot{Y}_3]$ .

Эта матрица, в частности, может быть получена из исходной матрицы СВП расчетной схемы замещения

$$[\dot{Y}] = \begin{bmatrix} \dot{Y}_{гг} & \dot{Y}_{гс} \\ \dot{Y}_{сг} & \dot{Y}_{сс} \end{bmatrix}$$

в соответствии с выражением

$$[\dot{Y}_3] = [\dot{Y}_{гг}] - [\dot{Y}_{гс}][\dot{Y}_{сс}]^{-1}[\dot{Y}_{сг}],$$

где индексом «с» обозначены СВП исключаемых (сетевых) узлов, а индексом «г» — СВП остающихся (генераторных) узлов с ЭДС.

Матрица  $W$ , на диагонали которой расположены собственные мощности генераторов ( $W_{ii}$ ), а остальные элементы ( $W_{ij} = W_{ji}$ ) представляют взаимные мощности между соответствующей парой генераторов, рассчитывается по формуле

$$W = E_{\text{diag}} Y_3' E_{\text{diag}},$$

где  $[E_{\text{diag}}]$  — диагональная матрица модулей ЭДС генераторов,  $[Y_3']$  — матрица, в которой диагональные элементы представляют собой активные составляющие собственных проводимостей ( $g_{ii}$ ) в матрице  $[\dot{Y}_3]$ , остальные элементы — модули взаимных проводимостей ( $y_{ij}$ ) в той же матрице.

Полученная матрица мощностей  $W$  является структурной моделью ЭЭС и может быть использована при анализе режимов и устойчивости ЭЭС.

В процессе оптимизации развития СЭС элементы матрицы  $W$  позволяют упрощенно определить предельные по условиям статической устойчивости мощности, передаваемые в сечениях электрической сети [20]. Сечением в электрической сети называется такая совокупность ее элементов, одновременное отключение которых разделяет систему на две изолированные части (подсистемы), в каждой из которых имеется хотя бы один генератор.

Сечения, для которых рассчитывается предельная по статической устойчивости мощность, могут быть условно разделены на два типа: генераторные (когда в одной подсистеме находится один генератор) и сетевые (все остальные).

Предельная по статической устойчивости мощность для первого типа сечений

$$P_q^{\text{пред}} = (1 - k_3) W_i, \quad (4)$$

где  $W_i = W_{ii} + \sum_{i \neq j} W_{ij}$  — максимальная мощность генератора, равная собственной структурной мощности и сумме взаимных структурных мощностей генератора  $i$  по всем его связям в эквивалентной схеме;  $k_3$  — коэффициент запаса по активной



мощности в сечении, который в соответствии с методическими указаниями [20] принимается равным 0,2.

Для второго типа сечений предельная по статической устойчивости мощность в  $q$ -м сечении, разделяющем систему на подсистемы  $A$  и  $B$  при передаче потока мощности из  $A$  и  $B$ :

$$P_q^{\text{пред}} = (1 - k_3) W_q^A, \quad (5)$$

где  $W_q^A = \sum_{i \in A} W_{ii} + \sum_{i \in A, j \neq A} W_{ij}$  — максимальная мощность сечения  $q$ , равная сумме собственных мощностей генераторов подсистемы  $A$  и сумме взаимных мощностей генераторов подсистем  $A$  и  $B$ , получаемая при условии

$$W_{ij} \geq \varepsilon_r P_{\text{ном}, i}^{\Gamma}, \quad (6)$$

где  $P_{\text{ном}, i}^{\Gamma}$  — номинальная мощность  $i$ -го генератора подсистемы  $A$ ,  $\varepsilon_r$  — малая величина, применяемая для оценки связности  $i$ -го генератора подсистемы  $A$  с генераторами подсистемы  $B$ .

Введение условия (6) позволяет рассчитать предельную по статической устойчивости мощность в сечении (4), исключив из нее собственные и взаимные мощности генераторов подсистемы  $A$ , имеющие слабые связи с подсистемой  $B$  (локальные или удаленные генераторы). В соответствии с работой [13] для генераторных узлов мощностью менее 3000 МВт значение  $\varepsilon_r$  рекомендуется принимать в интервале 0,1...0,15, а для узлов мощностью свыше 3000 МВт — в интервале 0,1...0,05.

## 2. МЕТОДИКА ОПТИМИЗАЦИИ РАЗВИТИЯ СИСТЕМООБРАЗУЮЩЕЙ ЭЛЕКТРИЧЕСКОЙ СЕТИ

В статической постановке предлагаемый методический подход к решению задачи оптимизации развития СЭС представлен на рис. 2. Методика включает в себя пять основных этапов.

На *I этапе* формируются исходные данные: граф электрической сети, отображающий существующие линии электропередачи (ЛЭП) и избыточный набор новых ЛЭП; располагаемые мощности станций  $P_i^{\text{ген}}$ ; мощности нагрузок в узлах  $P_i^{\text{потр}}$ ; удельные приведенные затраты на единицу передаваемой мощности по ЛЭП  $C_{ij}^{\text{лэп}}$ , включающие в себя капиталовложения и постоянные издержки, удельные переменные затраты на генерируемую мощность  $C_i^{\text{ген}}$ ; пропускные способности ЛЭП  $P_{ij}^{\text{лэп}}$ .

В качестве ограничения на пропускную способность существующих ЛЭП выступает наименьшая из предельных мощностей по нагреву провода или по статической устойчивости для отдельной связи с учетом коэффициента запаса.

Сформированные исходные данные используются на *II этапе* для получения решения на линейной модели оптимизации развития электрической сети, целевая функция которой

$$\min \sum_i \sum_j C_{ij}^{\text{лэп}} X_{ij}^{\text{лэп}} + C_i^{\text{ген}} X_i^{\text{ген}}, \quad (7)$$

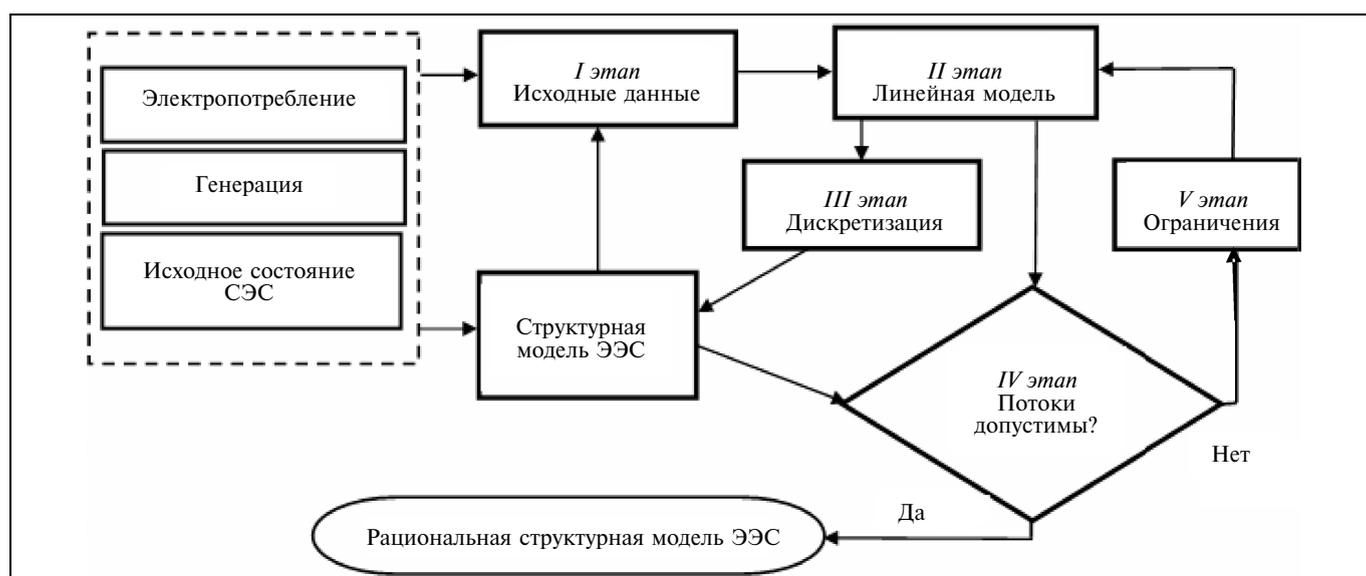


Рис. 2. Схема методики оптимизации развития системообразующей электрической сети

т. е. минимум приведенных затрат, при соблюдении балансов узлов ( $i = \overline{1, N}$ ):

$$X_i^{\text{ген}} + \sum_j (1 - p_{ij}) X_{ij}^{\text{ЛЭП}} - \sum_j X_{ij}^{\text{ЛЭП}} = P_i^{\text{потр}}, \quad (8)$$

и ограничений на потоки мощности по ЛЭП и рабочие мощности станций:

$$0 \leq X_{ij}^{\text{ЛЭП}} \leq P_{ij}^{\text{ЛЭП}}, \quad (9)$$

$$0 \leq X_i^{\text{ген}} \leq P_i^{\text{ген}}. \quad (10)$$

Неизвестными величинами являются потоки мощности по связям  $X_{ij}^{\text{ЛЭП}}$  из узла  $i$  в узел  $j$  и мощности генерации в узлах  $X_i^{\text{ген}}$ . В формуле (8)  $p_{ij}$  — удельный коэффициент потерь мощности при передаче по связи между узлами  $i$  и  $j$  (в относительных единицах).

Получаемое решение будет непрерывным — в виде минимально необходимых пропускных способностей межузловых связей электрической сети. В модели отсутствует алгоритм превращения этого непрерывного решения в дискретное (оптимальный вариант развития сети).

Поэтому *III этап* представляет собой дискретизацию решения, которая выполняется экспертом на основе данных о пропускной способности сечений СЭС, размере неиспользованной рабочей мощности станций в каждой из подсистем с учетом характеристик новых ЛЭП (номинального напряжения, пропускной способности, требуемых капиталовложений и др.). Эта процедура позволяет сформировать совокупность наиболее рациональных вариантов схем развития сети на основании значений потоков мощности по новым ЛЭП, а также учесть требования надежности в соответствии с требованиями [4]. Наличие потока, отличного от нуля, как правило, говорит о необходимости сооружения ЛЭП. При дискретизации решения определяется группа новых ЛЭП, которые

войдут в расчетную схему электрической сети, соответствующую одному из намеченных вариантов развития сети.

Для удобства графического представления сформированные варианты развития электрической сети отображаются на карте посредством модуля ГИС, связанного с программой «Сети», в которой реализован алгоритм рассматриваемой оптимизационной задачи. Этот модуль (рис. 3) входит в состав геоинформационного блока программно-вычислительного комплекса, содержащего также ряд специальных программ по геокодированию, обработке и выводу информации, необходимых для реализации ГИС-технологий [15].

В качестве ГИС-инструментария применяется пакет «MapInfo Professional». Этот программный продукт предоставляет для разрабатываемой системы достаточно графических возможностей и средств связывания объектов карты с атрибутивной информацией, адресного геокодирования, непосредственной работы с базой данных (БД), а также вывода на печать картографических материалов, их экспорта и импорта. Большое преимущество заключается в наличии прикладного языка программирования MapBasic. Если сравнивать «MapInfo» с другими ГИС-пакетами, то одним из лидеров по функциональным возможностям и использованию является пакет «ArcGIS» (целое семейство программных ГИС-продуктов ESRI). Эти ГИС-продукты целесообразно применять для расширенной обработки, анализа и моделирования геоданных с учетом топологических отношений [19]. В рамках настоящей работы базового набора функций пакета «MapInfo» достаточно для картографирования, визуализации и анализа данных, характеризующих СЭС.

На прикладном языке программирования MapBasic разработаны четыре программы — «GeoCode», «DrawUzel», «DrawLeps» и «SelVarT», позволяющие автоматизировать процессы картографирования узлов, ЛЭП и выбирать необходимый вариант расчета сети. Эти программы автоматически вызываются при работе геоинформационного модуля, первоначально загружающего пакет «MapInfo», который обрабатывает поступающие сообщения в фоновом режиме. Интерфейс модуля ГИС и пакета MapInfo осуществляется по технологии OLE Automation. После расчета варианта электрической сети посредством программы «Сети» при нажатии кнопки «ГИС» происходит географическое отображение линий электропередачи, стиль которых определяется программой «DrawLeps» на основе табличных данных, характеризующих тип линий, напряжение, наличие параллельных ЛЭП и другие признаки.

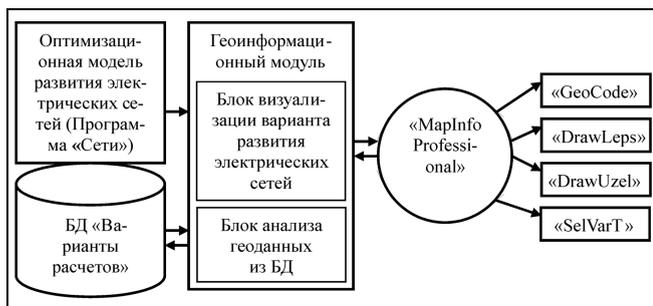


Рис. 3. Структура геоинформационного модуля

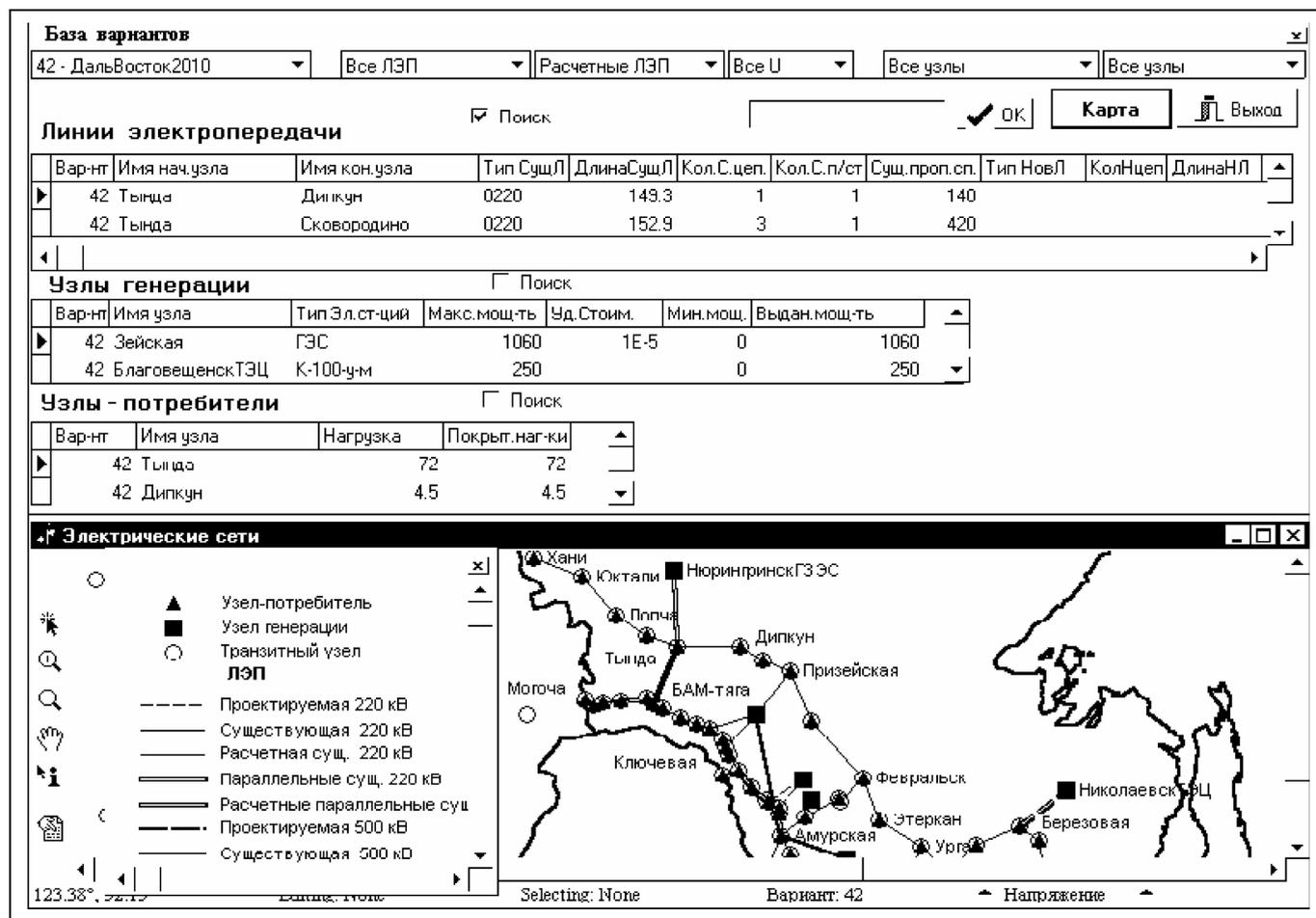


Рис. 4. Фрагмент работы геоинформационного модуля

Кроме того, имеется возможность непосредственной работы с БД «Варианты расчетов», когда вызываются необходимые таблицы из БД с соответствующей фильтрацией и пространственной привязкой данных. При этом в диалоговом окне «База вариантов» (программа «Сети») размещаются таблицы: «Варианты», «ЛЭП», «Генерация», «Потребители» и предоставляются соответствующие инструменты для выбора из таблиц требуемой информации (рис. 4).

Для выбранного варианта развития электрической сети возможны следующие тематические представления:

— ЛЭП — все ЛЭП; существующие ЛЭП; новые ЛЭП; ЛЭП с развитием;

— перетоки — все перетоки; линии с ненулевыми потоками; максимальный переток; минимальный переток;

— напряжение — все напряжения; 220; 330; 500; 750; 1150 кВ;

— генерация — все узлы генерации; узлы с максимальной мощностью; узлы с неиспользованной вырабатываемой мощностью;

— потребление — все узлы потребления; дефицитные узлы; самобалансирующиеся узлы; выдающие узлы; принимающие узлы (т. е. в зависимости от соотношения нагрузки и мощности генерирующих узлов).

Энергетические объекты наносятся на карту символами, специально созданными в соответствии с принятым стандартом графических обозначений электрических сетей ЭЭС. Геоинформационный модуль позволяет не только визуализировать варианты развития электрической сети, но и выполнить их анализ, используя средства графики, редактирования и обработки информации об энергетических объектах.

На IV этапе осуществляется проверка сформированных вариантов развития сети с точки зрения возможности передачи расчетных потоков мощности, включающая несколько стадий:

— определение контролируемых сечений и суммарного потока мощности в них по результатам расчета на линейной модели;

— составление расчетных схем замещения для сформированных вариантов электрической сети;

— расчет с помощью матрицы  $W$  предельно допустимых по статической устойчивости мощностей для контролируемых сечений (4) и (5);

— сравнение этих мощностей с полученными в линейной модели потоками мощности в тех же сечениях.

Если потоки мощности, рассчитанные по линейной модели, меньше допустимых для всех контролируемых сечений, то данный вариант схемы развития электрической сети считается обеспечивающим передачу расчетных потоков мощности, он включается в совокупность наилучших вариантов развития сети, подлежащих детальным проектным исследованиям на этапе их технико-экономического сравнения. В противном случае рассматриваемый вариант сети должен быть скорректирован (усилен).

Оптимальный способ усиления может быть получен посредством нового цикла оптимизации развития электрической сети, начиная с решения на линейной модели (II этап) с измененными значениями пропускных способностей сечений, в которых требуется усиление.

Для этого на  $V$  этапе модель (7)—(10) должна быть дополнена ограничениями вида

$$0 \leq \sum_{X_{ij}^{\text{лэп}} \in q} X_{ij}^{\text{лэп}} \leq P_q^{\text{пред}},$$

где  $P_q^{\text{пред}}$  — предельно допустимая по статической устойчивости мощность  $q$ -го сечения (4), (5).

После корректировки ограничений на потоки мощности в сечениях производится следующий цикл оптимизации развития электрической сети, т. е. осуществляется переход на II этап. На основе его результатов формируется новый вариант (группа вариантов) развития сети с увеличенными пропускными способностями сечений. Эта группа вариантов затем также должна быть проверена с помощью структурной модели ЭЭС. Итеративный процесс проверки и отбора совокупности наиболее рациональных вариантов развития сети продолжается до тех пор, пока все варианты развития сети, вошедшие в эту группу, не будут нуждаться в корректировке.

Учет фактора динамики (временного аспекта) развития электрической сети в работе предлагается осуществлять упрощенно — псевдодинамически. При этом последовательно рассматриваются

несколько этапов развития электрической сети, для каждого из которых применяется предложенная методика.

### 3. ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

Представленная методика и пакет программ оптимизации развития СЭС использовались при решении задачи развития электрической сети ОЭС Востока на среднесрочную перспективу [21], а также при разработке предложений по перспективам развития системообразующей электрической сети ЭЭС России на период до 2030 г.

Для проверки допустимости использования показателей структурного анализа для расчета пропускных способностей сечений (4) и (5) выполнены детализированные расчеты установившихся режимов и статической устойчивости с помощью специализированного программно-вычислительного комплекса АНАРЭС [22]. Эти расчеты подтвердили, что полученные с помощью предложенной методики схемы развития электрической сети удовлетворяют требованиям к эксплуатации оборудования и нормативам по статической устойчивости ЭЭС.

Вместе с тем, получаемые при использовании показателей структурного анализа значения пропускных способностей сечений отличаются от значений, полученных более точным методом [22], примерно на 15 % в сторону завышения. Как показали расчеты, такая погрешность при оптимизации развития электрической сети на период 15—20 лет допустима, поскольку в перспективе недостаток пропускной способности сечений может быть ликвидирован изменением состояния и режима существующих устройств регулирования напряжения в узлах (отключением шунтирующих реакторов, подключением статических компенсаторов, увеличением мощности синхронных компенсаторов и др.), а также вводом новых устройств. Кроме того, в редких случаях, когда дальнейшее увеличение числа новых линий не оказывает существенного влияния на прирост пропускной способности сечения, может потребоваться установка устройств продольной компенсации.

### ЗАКЛЮЧЕНИЕ: ВЫВОДЫ И РЕКОМЕНДАЦИИ

Предложен методический подход к решению задачи оптимизации развития системообразующей электрической сети. Его главная идея состоит в совместном использовании линейной модели оптимизации развития сети, структурной модели электроэнергосистемы и геоинформационных тех-



нологий в целях формирования совокупности наиболее рациональных вариантов развития системообразующей электрической сети, удовлетворяющих ограничениям на предельные по условиям статической устойчивости передаваемые мощности в сечениях электрической сети. Подход реализован в виде пакета программ, позволяющего более полно учитывать условия перспективного функционирования электрической сети.

Показана возможность использования при оптимизации развития электрической сети показателей структурного анализа для определения предельных мощностей в сечениях электроэнергосистемы. Полученные схемы развития электрической сети с помощью предложенной методики удовлетворяют требованиям к эксплуатации оборудования и нормативам по статической устойчивости системы.

Основные направления развития предложенного методического подхода и разработанных программных средств состоят в совершенствовании учета факторов надежности, многорежимности и динамики; учете особенностей развития электрической сети в условиях рынка; автоматизации формирования схем замещения электрической сети; интеграции линейной модели и структурной модели электроэнергосистемы в единый программный комплекс для анализа и оптимизации развития системообразующей электрической сети.

## ЛИТЕРАТУРА

1. *Системный* подход при управлении развитием электроэнергетики / Л.С. Беляев, Г.В. Войцеховская, В.А. Савельев и др. — Новосибирск: Наука, 1980.
2. *Постановление* Правительства Российской Федерации от 17 октября 2009 г. № 823 « О схемах и программах развития электроэнергетики».
3. *Справочник* по проектированию электроэнергетических систем / В.В. Ершевич, А.Н. Зейлигер, Г.А. Илларионов и др. — М.: Энергоатомиздат, 1985.
4. *Методические* рекомендации по проектированию развития энергосистем. — М.: Минэнерго России, 2003.
5. *Практические* рекомендации по оценке эффективности и разработке инвестиционных проектов и бизнес-планов в электроэнергетике (с типовыми примерами). Кн. 1. Методические особенности оценки эффективности проектов в электроэнергетике. — М.: НЦПИ, 2000.
6. *Лазебник А.И.* Применение метода ветвей и границ для выбора оптимальной электрической сети. — Изв. АН СССР. Энергетика и транспорт. — 1969. — № 2. — С. 138—143.
7. *Акишин Л.А., Макаревич А.Д., Молодюк В.В.* Статическая модель оптимизации конфигурации развивающейся электрической сети. — Тр. Иркутского политехнич. ин-та. — 1971. — Вып. 72. — С. 162—173.
8. *Дале В.А., Кришан З.П., Паэгле О.Г.* Оптимизация развития электрических сетей объединенных электроэнергетических систем методом динамического программирования. — Изв. АН СССР. Энергетика и транспорт. — 1970. — № 4. — С. 91—96.
9. *Моцкус И.Б.* О покоординатном методе оптимизации развития электрических сетей. — Изв. АН СССР. Энергетика и транспорт. — 1969. — № 1. — С. 54—60.
10. *Салливан Р.* Проектирование развития электроэнергетических систем. — М.: Энергоиздат, 1982.
11. *Попова О.М., Такайшвили В.Р., Труфанов В.В.* Пакет программ для анализа развития электрических сетей с использованием геоинформационных технологий: Препр. № 8 / ИСЭМ СО РАН. — Иркутск, 2001.
12. *Веников В.А.* Переходные электромеханические процессы в электрических системах. — М.: Энергия, 1964.
13. *Абраменкова Н.А., Вороний Н.И., Заславская Т.Б.* Структурный анализ электроэнергетических систем: В задачах моделирования и синтеза. — Новосибирск: Наука, 1990.
14. *Абраменкова Н.А., Вороний Н.И., Заславская Т.Б.* Структурно-режимный анализ электроэнергетических систем для выбора принципов противоаварийного управления. — Новосибирск: НГТУ, 1996.
15. *Попова О.М.* Построение геоинформационной системы электрических сетей // Вестник ИргТУ. — 2006. — № 2 (26). — С. 101—104.
16. *Геоинформатика* / А.Д. Иванников, В.П. Кулагин, А.Н. Тихонов, В.Я. Цветков. — М.: МАКС Пресс, 2001.
17. *Прикладная геоинформатика* / А.Д. Иванников, В.П. Кулагин, А.Н. Тихонов, В.Я. Цветков. — М.: МАКС Пресс, 2005.
18. *Цветков В.Я.* Геоинформационные системы и технологии. — М.: Финансы и статистика, 1998.
19. *Лурье И.К.* Геоинформационное картографирование. — М.: КДУ, 2010.
20. *Методические* указания по устойчивости энергосистем. — М.: ЦПТИ ОРГРЭС, 2004.
21. *Usov I.Yu.* A Linear Model and Structural Analysis for Main Grid Expansion Optimization // Liberalization and Modernization of Power Systems: Congestion Management Problems. The Intern. Workshop Proc. — Irkutsk: Energy System Institute, 2003. — P. 113—116.
22. *Программно-вычислительный* комплекс АНАРЭС—2000 для расчета и анализа нормальных и аварийных режимов ЭЭС / О.Н. Шепилов и др. // Современные программные средства для расчетов и оценивания состояния режимов ЭЭС: Сб. докладов II междунар. науч.-практ. семинара. — Новосибирск, ИДУЭС, 2002. — С. 5—14.

*Статья представлена к публикации членом редколлегии Б.В. Павловым.*

**Попова Ольга Михайловна** — канд. экон. наук, ст. науч. сотрудник, ☎ (3952) 42-63-80, ✉ POM@isem.sei.irk.ru,

**Усов Илья Юрьевич** — канд. техн. наук, мл. науч. сотрудник, ☎ (3952) 24-12-51, ✉ usov@govirk.ru,

Институт систем энергетики им. Л.А. Мелентьева СО РАН, г. Иркутск.

# ОПРЕДЕЛЕНИЕ КОЭФФИЦИЕНТОВ ВЛИЯНИЯ ПАРАМЕТРОВ РЕЖИМА ЭЛЕКТРОЭНЕРГЕТИЧЕСКОЙ СИСТЕМЫ НА ЕЕ УСТОЙЧИВОСТЬ

А.Е. Ефимов, О.Л. Опалев, И.Б. Ядыкин

Выполнен анализ устойчивости электроэнергетической системы путем исследования свойств матрицы динамики ее линеаризованной модели. Введена мера устойчивости, зависящая от параметров электроэнергетической системы, решена задача оценивания запаса устойчивости ее режимов. Разработан алгоритм вычисления коэффициентов влияния параметров режима на показатель устойчивости.

**Ключевые слова:** электроэнергетическая система, математическая модель, устойчивость по Ляпунову, запас устойчивости, матрица динамики, показатель устойчивости, коэффициент влияния.

## ВВЕДЕНИЕ

Электроэнергетическая система (ЭЭС), обеспечивающая работу промышленности, транспорта, быт населения — всю жизнедеятельность охваченных электрической сетью территорий, должна работать надежно, прежде всего, устойчиво [1]. Под устойчивой работой ЭЭС (или, просто, устойчивостью ЭЭС) понимается синхронность работы ее генераторов или, в терминах режимного управления, постоянство относительных фазовых углов их ЭДС.

Исследование устойчивости ЭЭС базируется на математической модели в виде системы дифференциальных уравнений, описывающей ее работу и взаимодействие ее элементов. В данной статье под устойчивостью ЭЭС понимается устойчивость по Ляпунову, а для ее исследования применяется метод первого приближения и теоремы Ляпунова.

При анализе устойчивости ЭЭС важно оценить коэффициенты влияния параметров режима на устойчивость. Под параметрами режима понимаются показатели, количественно определяющие условия работы системы. К ним относятся амплитуды и фазы напряжений, токов, ЭДС, а также активные и реактивные мощности и др.

## 2. ОПРЕДЕЛЕНИЕ ПОКАЗАТЕЛЯ УСТОЙЧИВОСТИ

### 2.1. Описание электроэнергетической системы. Матрица динамики системы

В качестве математической модели ЭЭС рассматривается система дифференциальных уравнений движения роторов генераторов без учета демпфирования, независимыми переменными служат абсолютные и относительные фазовые углы ЭДС генераторов:

$$\begin{aligned} T_1 \frac{d^2 \delta_1}{dt^2} &= P_{10} - P_1(\delta_{12}, \delta_{13}, \dots, \delta_{1n}), \\ &\dots \\ T_n \frac{d^2 \delta_n}{dt^2} &= P_{n0} - P_n(\delta_{12}, \delta_{13}, \dots, \delta_{1n}), \end{aligned} \quad (1)$$

где  $P_{i0}$  — механическая мощность на валу  $i$ -го генератора;  $i = 1, \dots, n$ ;  $P_i$  — электрическая мощность, отдаваемая в сеть  $i$ -м генератором;  $\delta_i$  — фаза ЭДС  $i$ -го генератора;  $\delta_{1i} = \delta_1 - \delta_i$  — относительный фазовый угол между фазовыми углами ЭДС первого и  $i$ -го генераторов;  $T_i$  — постоянная инерции ротора генератора;  $n$  — число генераторов. Предполагается, что изменение скорости вращения ротора отдельного генератора на протяжении всего



времени рассматриваемого процесса значительно меньше, чем синхронная скорость [1].

Вид системы уравнений (1) позволяет привести ее к нормальной форме Коши и применить метод первого приближения и теоремы Ляпунова об устойчивости.

После линеаризации в окрестности некоторой исходной стационарной точки  $\{\delta_i(0), \dots, \delta_n(0)\}$  система уравнений (1) примет вид

$$\begin{aligned} T_1 \frac{d^2 \Delta \delta_1}{dt^2} + \frac{\partial P_1}{\partial \delta_{12}} \Delta \delta_{12} + \dots + \frac{\partial P_1}{\partial \delta_{1n}} \Delta \delta_{1n} &= 0, \\ T_n \frac{d^2 \Delta \delta_n}{dt^2} + \frac{\partial P_n}{\partial \delta_{12}} \Delta \delta_{12} + \dots + \frac{\partial P_n}{\partial \delta_{1n}} \Delta \delta_{1n} &= 0, \end{aligned} \quad (2)$$

где  $\Delta \delta_i = \delta_i(t) - \delta_i(0)$ .

Деление каждого из уравнений системы (2) на соответствующее значение  $T_i$  и вычитание каждого уравнения из первого позволяют привести ее к виду

$$\begin{aligned} \frac{d^2 \Delta \delta_{12}}{dt^2} + \left( \frac{1}{T_1} \frac{\partial P_1}{\partial \delta_{12}} - \frac{1}{T_1} \frac{\partial P_1}{\partial \delta_{12}} \right) \Delta \delta_{12} + \dots \\ \dots + \left( \frac{1}{T_1} \frac{\partial P_1}{\partial \delta_{1n}} - \frac{1}{T_2} \frac{\partial P_2}{\partial \delta_{1n}} \right) \Delta \delta_{1n} &= 0, \\ \frac{d^2 \Delta \delta_{1n}}{dt^2} + \left( \frac{1}{T_1} \frac{\partial P_1}{\partial \delta_{12}} - \frac{1}{T_n} \frac{\partial P_n}{\partial \delta_{12}} \right) \Delta \delta_{12} + \dots \\ \dots + \left( \frac{1}{T_1} \frac{\partial P_1}{\partial \delta_{1n}} - \frac{1}{T_n} \frac{\partial P_n}{\partial \delta_{1n}} \right) \Delta \delta_{1n} &= 0, \end{aligned} \quad (3)$$

или в матричной форме  $\dot{X} + AX = 0$ , где  $X = (\Delta \delta_{12}, \Delta \delta_{13}, \dots, \Delta \delta_{1n})^T$ .

Путем замены  $\dot{X} = Y$  система приводится к виду:

$$\begin{vmatrix} \dot{Y} \\ \dot{X} \end{vmatrix} = \begin{vmatrix} 0 & -A \\ E & 0 \end{vmatrix} \cdot \begin{vmatrix} Y \\ X \end{vmatrix}. \quad (4)$$

Матрица данной системы дифференциальных уравнений является *матрицей динамики* энергосистемы.

## 2.2. Показатель устойчивости

Система дифференциальных уравнений (4) — линейная однородная система с постоянными коэффициентами, записанная в нормальной форме Коши. Если задать начальное условие, то решение задачи Коши можно найти аналитически. При этом характер движения системы полностью определяется собственными числами матрицы коэф-

фициентов дифференциальных уравнений, т. е. матрицы динамики энергосистемы.

Ответ на вопрос устойчива или неустойчива энергосистема, описываемая системой уравнений (4), дают две теоремы Ляпунова [2], согласно которым степень устойчивости определяется собственным числом с максимальной действительной частью.

Назовем действительную часть такого собственного числа *показателем устойчивости* системы (4) и обозначим его  $\chi$ :

$$\chi = \max_k (\operatorname{Re}(\lambda_k)) = \operatorname{Re}(\lambda_\chi), \quad k = \overline{1, 2n-2},$$

где  $\lambda_k$  — собственное число матрицы динамики,  $2n-2$  — ее ранг.

Анализ аналитического решения [3] системы (4) показывает, что если показатель устойчивости положителен, то ЭЭС неустойчива и, следовательно, при некотором малом начальном возмущении теряет устойчивость.

Поскольку ненулевые собственные числа матриц систем (3) и (4) совпадают, то показатели устойчивости данных систем также совпадают. Следовательно, для изучения устойчивости ЭЭС достаточно ограничиться рассмотрением матрицы системы (3) — матрицы  $A$ . Далее по тексту под матрицей динамики подразумевается  $A$ .

## 3. ОПРЕДЕЛЕНИЕ КОЭФФИЦИЕНТОВ ВЛИЯНИЯ

Для определения коэффициента влияния параметров режима на устойчивость энергосистемы рассмотрим произвольный стационарный режим ее работы, описываемый системой уравнений (3) с соответствующим показателем устойчивости. При малом изменении какого-либо одного параметра режима исходный режим перейдет в другой, близкий ему и тоже стационарный. Показатель устойчивости нового режима при этом будет отличаться от исходного показателя устойчивости на некую малую величину. Таким образом, можно определить коэффициент влияния данного параметра режима как коэффициент пропорциональности между малым изменением параметра режима и соответствующим малым изменением показателя устойчивости энергосистемы.

Другими словами коэффициент влияния параметра режима на устойчивость энергосистемы есть частная производная показателя устойчивости по данному параметру при квазистационарном изменении режима:  $k_{p_g} = \partial \chi / \partial p_g$ , где  $p_g$  — параметр режима энергосистемы.

В соответствии с правилом дифференцирования сложной функции

$$k_{p_g} = \sum_{i,j=1}^{n-1} \frac{\partial \chi}{\partial a_{ij}} \cdot \frac{\partial a_{ij}}{\partial p_g}, \quad (5)$$

где  $a_{ij}$  — элемент матрицы  $A$  динамики энергосистемы ранга  $n - 1$  ( $n$  — число генераторов системы) с показателем устойчивости  $\chi$ .

### 3.1. Лемма о вычислении частных производных собственного числа матрицы динамики по элементам матрицы

Выражение (5) для расчета коэффициента влияния параметра режима включает в себя частные производные действительной части собственного числа матрицы динамики  $A$  по элементам этой матрицы  $a_{ij}$ . Так как все элементы матрицы динамики числа действительные, то для данных частных производных справедливо

$$\frac{\partial \operatorname{Re}(\lambda_k)}{\partial a_{ij}} = \operatorname{Re} \left( \frac{\partial \lambda_k}{\partial a_{ij}} \right).$$

Для вычисления частной производной в правой части данного равенства сформулируем следующую лемму.

**Лемма.** Если  $\lambda_k$  — собственное число матрицы  $A = \|a_{ij}\|$ ,

$$\frac{\partial \lambda_k}{\partial a_{ij}} = \frac{x_{ji}}{\operatorname{tr}(X)}, \quad (6)$$

где  $X$  — присоединенная (союзная) матрица к характеристической матрице матрицы  $A$ :  $X = \operatorname{adj}(A - \lambda_k E)$ . ♦

Отметим, что частные производные собственных чисел по элементам матрицы могут иметь точки разрыва. Однако они находятся за областью существования режимов физической системы, описываемой представленной в статье моделью, и поэтому здесь не рассматриваются.

Доказательство леммы приведено в Приложении.

### 3.2. Расчет элементов матрицы динамики

Следующий шаг расчета коэффициентов влияния состоит в определении элементов матрицы динамики энергосистемы. Для этого необходимо выразить активную мощность, отдаваемую в сеть  $k$ -м генератором, через относительные фазовые углы ЭДС. Такую зависимость принято называть характеристикой мощности генератора [1, 4].

Для расчета электрического тока через  $k$ -й генератор применяется метод наложения, в соответствии с которым

$$I_k = E_1 Y_{k1} + E_2 Y_{k2} + \dots + E_n Y_{kn} = \sum_{i=1}^n E_i Y_{ki}.$$

Собственные и взаимные проводимости электрической системы ( $Y_{ii}$  и  $Y_{ki}$ ) находятся путем комбинации метода единичных ЭДС [1] и решения системы уравнений баланса мощностей.

Комплексное значение мощности, отдаваемой генератором в сеть, определяется как произведение комплексного напряжения, действующего на шинах генератора, на сопряженный комплексный ток:

$$S_k = U_k \bar{I}_k = U_k \overline{E_1 Y_{k1}} + U_k \overline{E_2 Y_{k2}} + \dots \\ \dots + U_k \overline{E_n Y_{kn}} = U_k \sum_{i=1}^n \overline{E_i Y_{ki}}.$$

Если перейти к экспоненциальной форме записи  $U_1 = U_1 e^{j\gamma_1}$ ,  $E_k = E_k e^{j\delta_k}$ ,  $Y_{pq} = Y_{pq} e^{j\psi_{pq}}$ , то

$$S_k = U_k \sum_{i=1}^n \overline{E_i Y_{ki}} = U_k e^{j\gamma_k} \sum_{i=1}^n E_i e^{-j\delta_i} Y_{ki} e^{j\psi_{ki}} = \\ = \sum_{i=1}^n U_k E_i Y_{ki} e^{j(\gamma_k - \delta_i - \psi_{ki})}.$$

Соответственно, отдаваемая активная мощность генератора

$$P_k = \sum_{i=1}^n U_k E_i Y_{ki} \cos(\gamma_k - \delta_i - \psi_{ki}).$$

После замены аргументов  $\psi_{pq}$  дополняющими их до  $90^\circ$  аргументами  $\alpha_{pq} = 90^\circ - \psi_{pq}$  и введения разности фаз между напряжением и ЭДС  $k$ -го генератора ( $\delta_k^0 = \gamma_k - \delta_k$ ) и разности фаз между ЭДС  $k$ -го и  $i$ -го генераторов ( $\delta_{ki} = \delta_k - \delta_i$ ) характеристика мощности в относительных фазовых углах принимает вид

$$P_k(\delta_{11}, \dots, \delta_{1n}) = \\ = \sum_{i=1}^n U_k E_i Y_{ki} \sin(\delta_k^0 + (\delta_{1i} - \delta_{1k}) + \alpha_{ki}). \quad (7)$$

Расчет частных производных полученных функций мощностей генераторов позволяет получить



в аналитическом виде элементы матрицы динамики энергосистемы (см. систему (3)):

$$\begin{aligned}
 a_{j,k} &= \frac{1}{T_1} \frac{\partial P_1}{\partial \delta_{1,k+1}} - \frac{1}{T_{j+1}} \frac{\partial P_{j+1}}{\partial \delta_{1,k+1}} = \\
 &= \frac{1}{T_1} U_1 E_{k+1} Y_{1,k+1} \cos(\delta_1^0 + \delta_{1,k+1}) - \\
 &- \frac{1}{T_{j+1}} U_{j+1} E_{k+1} Y_{j+1,k+1} \cos(\delta_{j+1}^0 + \\
 &+ (\delta_{1,k+1} - \delta_{1,j+1}) + \alpha_{j+1,k+1}) + \\
 &+ \delta(j,k) \frac{1}{T_{j+1}} \sum_{i=1}^n U_{j+1} E_i Y_{j+1,i} \cos(\delta_{j+1}^0 + \\
 &+ (\delta_{1i} - \delta_{1,j+1}) + \alpha_{j+1,i}), \quad (8)
 \end{aligned}$$

где  $\delta(j,k)$  — символ Кронекера,  $T_i$  — постоянная инерции  $i$ -го генератора,  $P_i$  — активная мощность  $i$ -го генератора,  $U_i$  — напряжение в узле, к которому присоединен  $i$ -й генератор,  $E_i$  — ЭДС  $i$ -го генератора,  $Y_{jk}$  — элемент матрицы проводимостей,  $\alpha_{jk}$  — величина, дополняющая до  $90^\circ$  фазовый угол проводимости  $Y_{jk}$ ,  $\delta_i^0$  — угол между напряжением и ЭДС  $i$ -го генератора,  $\delta_{1i} = \delta_1 - \delta_i$  — относительный угол между ЭДС 1-го и  $i$ -го генераторов,  $n$  — число генераторов.

### 3.3. Алгоритм поиска коэффициентов влияния

Алгоритм поиска коэффициентов влияния параметров  $p_g$  режима на устойчивость электрической энергосистемы состоит из следующих основных шагов.

**Шаг 1.** Определение базиса независимых переменных, однозначно описывающих режим энергосистемы при известных системных параметрах, таких как проводимости линий, собственные сопротивления генераторов и др. Например, в качестве такого базиса можно использовать амплитуды и фазы ЭДС всех генераторов системы либо активные и реактивные мощности генераторов:

$$\begin{aligned}
 b &= \{E_1, \dots, E_n, \delta_1, \dots, \delta_n\} \text{ или} \\
 b &= \{P_1, \dots, P_n, Q_1, \dots, Q_n\}.
 \end{aligned}$$

**Шаг 2.** Вычисление частных производных параметров режима, входящих в характеристики мощности генераторов (7), и параметров  $p_g$  режима по выбранному базисным переменным. Для этого необходимо записать уравнения балансов активных и реактивных мощностей в узлах ЭЭС (уравнения установившегося режима [5]) в базисных переменных и выразить из них искомые частные производные  $\partial U_k / \partial b_l$ ,  $\partial \delta_k^0 / \partial b_l$ ,  $\partial E_k / \partial b_l$ ,  $\partial \delta_{1k}^0 / \partial b_l$ ,  $\partial p_g / \partial b_l$ .

**Шаг 3.** Вычисление частных производных элементов матрицы динамики по базисным переменным в соответствии с правилом дифференцирования сложной функции

$$\begin{aligned}
 \frac{\partial a_{ij}}{\partial b_l} &= \sum_k \left( \frac{\partial a_{ij}}{\partial U_k} \frac{\partial U_k}{\partial b_l} + \frac{\partial a_{ij}}{\partial E_k} \frac{\partial E_k}{\partial b_l} + \right. \\
 &+ \left. \frac{\partial a_{ij}}{\partial \delta_k^0} \frac{\partial \delta_k^0}{\partial b_l} + \frac{\partial a_{ij}}{\partial \delta_{1k}^0} \frac{\partial \delta_{1k}^0}{\partial b_l} \right), \quad (9)
 \end{aligned}$$

где частные производные  $\frac{\partial a_{ij}}{\partial U_k}$ ,  $\frac{\partial a_{ij}}{\partial E_k}$ ,  $\frac{\partial a_{ij}}{\partial \delta_k^0}$ ,  $\frac{\partial a_{ij}}{\partial \delta_{1k}^0}$  рассчитываются из выражения (8).

**Шаг 4.** Вычисление частных производных элементов матрицы динамики по параметрам  $p_g$  режима:

$$\frac{\partial a_{ij}}{\partial p_g} = \sum_{i=1}^{2n} \frac{\partial a_{ij}}{\partial b_l} \left( \frac{\partial p_g}{\partial b_l} \right)^{-1}.$$

**Шаг 5.** Определение коэффициентов влияния

$$\begin{aligned}
 k_{p_g} &= \frac{\partial \chi}{\partial p_g} = \sum_{i,j=1}^{n-1} \frac{\partial \chi}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial p_g} = \\
 &= \sum_{i,j=1}^{n-1} \operatorname{Re} \left( \frac{\partial \lambda_{\chi}}{\partial a_{ij}} \right) \sum_{l=1}^{2n} \frac{\partial a_{ij}}{\partial b_l} \left( \frac{\partial p_g}{\partial b_l} \right)^{-1}, \quad (10)
 \end{aligned}$$

где  $\lambda_{\chi}$  — собственное число матрицы динамики с максимальной действительной частью.

Итак, коэффициенты влияния параметров режима на устойчивость энергосистемы вычисляются по выражению (10) через частные производные (6) собственного числа матрицы динамики по элементам матрицы (8) и частные производные элементов матрицы динамики по базисным переменным (9).

Ранжирование рассчитываемых значений коэффициентов влияния при нахождении показателя устойчивости в зоне минимального запаса устойчивости или переходе в область, соответствующую неустойчивому состоянию системы, позволит идентифицировать наиболее проблемные элементы энергосистемы.

## ЗАКЛЮЧЕНИЕ

Описанная методика оценивания устойчивости режимов электроэнергетических систем основана на классической теории устойчивости по Ляпунову и представлении энергосистемы в виде физико-математической модели. На основе теорем Ляпунова предложена мера устойчивости энергосистемы,

основанная на вычислении спектра матрицы динамики системы и выделении корня характеристического уравнения с максимальным значением вещественной части. Для построения эффективных законов управления энергосистемой важно определить те параметры режима, которые обеспечивают наибольший отклик на приложенное изменение параметра режима.

Предложенный алгоритм расчета коэффициентов влияния параметров режима на устойчивость электроэнергетической системы может быть применен для определения оптимальных управляющих воздействий, позволяющих выйти из зоны опасной близости к предельному режиму, либо для управления коэффициентом запаса энергосистемы по устойчивости. Ранжирование значений коэффициентов влияния позволит выделить наиболее проблемные в смысле влияния на устойчивость элементы энергосистемы.

При расчете коэффициентов влияния параметров режима на устойчивость электроэнергетической системы важную роль играет лемма о вычислении частных производных собственного числа матрицы по ее элементам. Ввиду большой размерности матрицы динамики энергетической системы, любое решение задачи анализа устойчивости энергосистемы включает в себя решение задачи редукции математической модели.

Опираясь на описанную методику и результаты анализа влияния параметров режима на устойчивость электроэнергетической системы, можно сформировать эквивалентные модели ЭЭС, удобные для анализа ее устойчивости.

#### ПРИЛОЖЕНИЕ

Доказательство леммы. Обозначим  $B = A - \lambda E$  и запишем характеристическое уравнение  $\det(B) = 0$ , из которого следует  $d\det(B) = 0$ .

Согласно формуле Якоби [6]  $d\det(B) = \text{tr}(\text{adj}(B)dB)$ .

Из равенства дифференциала детерминанта нулю следует, что

$$\text{tr}(\text{adj}(B)dB) = \text{tr}(\text{adj}(A - \lambda E)(dA - d\lambda E)) = 0.$$

Обозначим  $X = \text{adj}(A - \lambda E)$ , тогда

$$\text{tr}(X(dA - d\lambda E)) = \text{tr}(XdA) - d\lambda \text{tr}(X) = 0. \quad (11)$$

$$\text{Если } da_{ij} = \begin{cases} 0, & i \neq r, j \neq s, \\ da_{rs}, & i = r, j = s, \end{cases}$$

то

$$(XdA)_{ij} = \sum_{k=1}^n x_{ik} da_{kj} = \begin{cases} 0, & j \neq s, \\ x_{ir} da_{rs}, & j = s. \end{cases}$$

Соответственно, в матрице  $(XdA)$  только в  $s$ -м столбце имеются элементы, отличные от нуля. Из этого следует

$$\text{tr}(XdA) = (XdA)_{s,s} = x_{sr} da_{rs}.$$

Подставив данное выражение в формулу (11), получим

$$x_{sr} da_{rs} - d\lambda \text{tr}(X) = 0.$$

Следовательно

$$\frac{\partial \lambda}{\partial a_{rs}} = \frac{x_{sr}}{\text{tr}(X)}. \quad \blacklozenge$$

#### ЛИТЕРАТУРА

1. *Веников В.А.* Переходные электромеханические процессы в электрических системах. — М.: Высшая школа, 1985. — 536 с.
2. *Методы классической и современной теории управления* // Под ред. К.А. Пупкова и Н.Д. Егупова — М.: Изд-во МГТУ им. Н.Э. Баумана, 2004.
3. *Давыдов В.В., Неуймин В.Г., Сактоев В.Е.* Определение критических сечений энергосистем в предельных режимах // Российская академия наук. Энергетика и транспорт. — Улан-Удэ: 1992. — Том 38.
4. *Куликов Ю.А.* Переходные процессы в электрических системах. — Новосибирск: Изд-во НГТУ, 2002. — 283 с.
5. *Идельчик В.И.* Расчеты установившихся режимов электрических систем. — М.: Энергия, 1977. — 192 с.
6. *Magnus J.R., Neudecker H.* Matrix Differential Calculus with Applications in Statistics and Econometrics. — N.Y.: John Wiley & Sons, 1999. — 395 с.

*Статья представлена к публикации членом редколлегии В.Ю. Рутковским.*

**Ефимов Артем Евгеньевич** — аналитик Департамента по развитию, ЗАО «Институт энергетических систем»,  
✉ efimov.artem@gmail.com,

**Опалев Олег Леонидович** — вед. специалист,  
ОАО «Системный оператор Единой энергетической системы»,  
✉ opalev\_oleg@mail.ru,

**Ядыкин Игорь Борисович** — д-р техн. наук, зав. лабораторией,  
Институт проблем управления им. В.А.Трапезникова РАН,  
☎ (495) 334-90-31, ✉ Jad@ipu.ru.



# РАСЧЕТ АППРОКСИМИРУЮЩИХ ФУНКЦИЙ ДВУМЕРНЫХ ТЕМПЕРАТУРНЫХ ПОЛЕЙ В ЗАДАЧАХ УПРАВЛЕНИЯ ПОДВИЖНЫМИ ИСТОЧНИКАМИ ВОЗДЕЙСТВИЙ

В.И. Финягина

Метод подстановки для вычисления подвижных управляющих воздействий в системах с распределенными параметрами распространен на двумерный случай. Рассмотрена задача подвижного управления, состоящая в получении и поддержании в двумерной области заданного поля температур. Предложено аппроксимировать конечными суммами ряда Фурье заданное в виде кусочно-постоянных функций поле температур. Показано, что такая аппроксимация позволяет получить достаточно крутые фронты изменения температуры при переходе от одного участка постоянства температуры к другому. Приведены результаты численного моделирования.

**Ключевые слова:** система, распределенные параметры, подвижное управление, метод подстановки, двумерное температурное поле, частичная сумма ряда Фурье.

## ВВЕДЕНИЕ

Рассматриваемая в настоящей статье задача относится к классу задач управления системами с распределенными параметрами с помощью подвижных управляющих воздействий [1]. Конкретно, рассматриваются объекты, состояние которых описывается двумерным уравнением теплопроводности, источник тепла перемещается вдоль определенной траектории в двумерной области обрабатываемого изделия. Задача подвижного управления состоит в получении и поддержании в этой области распределения температур, близких к заданным.

Последовательно применяя предложенные ранее *метод подстановки* для расчета распределенных управлений [2] и *метод реализации* распределенных управлений с помощью подвижных источников воздействия [3], можно найти траекторию и закон движения такого подвижного управления, которое, будучи приложено к объекту управления, переводит его из некоторого заданного состояния в требуемое. Однако оба эти метода разработаны

для объектов, состояние которых описывается одномерным уравнением теплопроводности.

В работе [4] представлен метод подстановки для расчета распределенных управлений в двумерной области. Основную трудность при переходе к двумерной задаче представляет расчет функции, аппроксимирующей требуемое распределение температуры на поверхности объекта. В упомянутой работе требуемое распределение температуры аппроксимировалось квадратичными функциями.

В данной работе функция двух переменных аппроксимируется частичной (конечной) суммой ряда Фурье. Необходимость применения предлагаемого здесь способа диктуется условиями задачи. Если на такой же площади поверхности обрабатываемого объекта необходимо создать несколько областей с различной температурой, которые могут располагаться достаточно близко друг к другу и даже могут соприкасаться, то предлагаемый далее способ может оказаться предпочтительнее, поскольку дает возможность получить довольно крутой фронт изменения температуры при переходе от одной области к другой и исключает необходимость создания так называемых областей «сшивки».

## 1. МЕТОД ПОДСТАНОВКИ ДЛЯ РАСЧЕТА РАСПРЕДЕЛЕННЫХ УПРАВЛЕНИЙ ОТ ТЕПЛОВЫХ ИСТОЧНИКОВ, ДВИЖУЩИХСЯ ПО ПОВЕРХНОСТИ

Напомним коротко (поскольку в работе [4] это сделано достаточно подробно) суть метода подстановки применительно к двумерному объекту.

Состояние объекта определяется функцией  $Q(x, t)$ ,  $x = (x_1, x_2) \in D$ ,  $t \geq t_0$ , которая описывает распределение температуры в двумерной области  $D$  и удовлетворяет (вообще говоря, нелинейному) уравнению теплопроводности:

$$c\rho \frac{\partial Q}{\partial t} - L(Q) + \varphi(Q) = f(x, t), \\ x = (x_1, x_2) \in D, \quad t \geq 0, \quad (1.1)$$

с граничными и начальными условиями

$$\left[ \alpha Q + \lambda \frac{\partial Q}{\partial n} \right]_{x \in \Gamma} = \alpha Q_\Gamma, \quad \Gamma = \partial D, \\ Q(x, 0) = Q_0(x), \quad x \in D. \quad (1.2)$$

Здесь оператор  $L(Q)$ , действующий по пространственной переменной  $x$  на функцию  $Q = Q_0(x, t)$ , определяется равенством

$$L(Q) = \operatorname{div}[\lambda \operatorname{grad} Q]. \quad (1.3)$$

В приведенных выражениях  $t$  — время,  $x = (x_1, x_2) \in R^2$ ;  $D$  — ограниченная область на плоскости;  $c$ ,  $\rho$ ,  $\lambda$  и  $\alpha$  — теплотехнические параметры материала объекта, соответственно теплоемкость, плотность, теплопроводность и коэффициент теплообмена между поверхностью и внешней средой, зависящие в общем случае от местоположения (координат  $x_1, x_2$ ) точки  $x$  и температуры  $Q$ ;  $\varphi(Q)$  — функция, описывающая теплоотвод с поверхности объекта во внешнюю среду;  $\frac{\partial Q}{\partial n}$  — производная по направлению внешней нормали к границе  $\Gamma$  области  $D$ . Функция  $f(x, t)$  описывает воздействие от подвижного источника тепла с параметрами интенсивность, форма и, возможно, другими параметрами, например, дисперсии.

При построении аппроксимирующей функции температурного поля рассматривается *установившийся* режим — это *многоцикловый* режим движения источника [5, 6]. В статье [5] на основании теоремы Н.Н. Боголюбова об усреднении доказано, что при малом периоде  $T$  (большой скорости движения источника) в установившемся режиме  $\frac{\partial Q(x, t)}{\partial t} = 0$ ,  $f(x, t) \approx f(t)$  (тепловое воздействие реализуется с достаточной степенью точности), оператор (1.3) (при условии  $\lambda = \operatorname{const}$ )  $L(Q) =$

и уравнение объекта (1.1) упрощается и принимает вид

$$-L(Q) + \varphi(Q) = f(x), \quad x = (x_1, x_2) \in D. \quad (1.4)$$

Обозначим желаемое состояние объекта  $Q^*(x)$ , а аппроксимирующую функцию  $Q_A^*(x)$  и вместо уравнения (1.1) получим следующий вид искомого распределенного воздействия:

$$f^*(x) = -L[Q_A^*(x)] + \varphi(Q_A^*). \quad (1.5)$$

Аппроксимирующая функция  $Q_A^*(x)$  должна:

- 1) удовлетворять граничным условиям (1.2);
- 2) быть кусочно-дифференцируемой такое число раз, которое необходимо для подстановки в оператор  $L$ ;
- 3) аппроксимировать желаемое состояние  $Q^*(x)$  в области  $D_1$ ,  $D_1 \subset D$ , с требуемой точностью.

Функция (1.5) должна быть кусочно-непрерывной и удовлетворять в области  $D$  условиям:  $f^*(x) \geq 0$ ,  $\int_D f^*(x) dx \leq U_{\max}$  ( $U_{\max}$  — максимальная

интенсивность (мощность) источника) и быть равной нулю вне области  $D$ .

Требования 1 и 2 необходимы для того чтобы функция  $Q_A^*(x)$  удовлетворяла уравнению (1.4) с граничными условиями (1.2). Это и делает возможным расчет воздействия  $f^*(x)$  по формуле (1.5).

Рассмотрим аппроксимацию разрывной функции частичной (конечной) суммой ряда Фурье, которую будем называть полиномом Фурье [7]. При таком способе аппроксимации можно добиться довольно крутого фронта возрастания температуры в области перехода от нулевых граничных условий к значению заданного распределения  $Q^*(x)$ .

## 2. РАСЧЕТ АППРОКСИМИРУЮЩЕЙ ФУНКЦИИ И РАСПРЕДЕЛЕННОГО УПРАВЛЕНИЯ С ПОМОЩЬЮ ПОЛИНОМОВ ФУРЬЕ

Пусть объект представляет собой плоскую ограниченную тонкую пластину с размерами  $a$  и  $b$  по осям  $x$  и  $y$  соответственно (для избавления от «лишних» индексов положим  $x_1 = x$  и  $x_2 = y$ ) и толщиной  $h$ ,  $h \ll a$ ,  $h \ll b$ . Распределением температуры по толщине пренебрегаем. Пусть область  $D$  совпадает с областью  $D_1$  и имеет вид прямоугольника со сторонами  $a$  и  $b$ . Нелинейный член уравнения (1.1)  $\varphi(Q)$  описывает теплоотдачу с поверхности объекта по закону Стефана—Больцмана  $\varphi(Q) = \varepsilon\sigma(Q^4 - Q_o^4)$  ( $Q_o$  — температура окружающей среды). Желаемое состояние объекта  $Q^*(x, y)$  задано в виде кусочно-постоянной функции, а



$$= \lambda \left( \frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2} \right).$$

С учетом сказанного распределенное воздействие (1.5)

$$f^*(x, y) = -\lambda \left( \frac{\partial^2 Q}{\partial x^2} + \frac{\partial^2 Q}{\partial y^2} \right) + \varepsilon \sigma (Q^{*4} - Q_o^4). \quad (2.1)$$

Будем считать, что температурное поле  $Q^*(x, y)$  в установившемся режиме равномерное, а переход от нулевых граничных условий к значению  $Q^*(x, y)$  есть прямые линии разрыва, которые пересекают ось  $x$  в точках  $x = 0, x = a$ , а ось  $y$  — в точках  $y = 0, y = b$ .

Запишем аппроксимирующую функцию в виде следующего полинома Фурье:  $Q_A^*(x, y) =$

$$= \sum_{n=1}^N \sum_{m=1}^M C_n \sin \frac{n\pi x}{a} C_m \sin \frac{m\pi y}{b} \quad \text{или, обозначив}$$

$$A_{nm} = C_n C_m,$$

$$Q_A^*(x, y) = \sum_{n=1}^N \sum_{m=1}^M A_{nm} \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{b}. \quad (2.2)$$

Сделав расчеты, необходимые для подстановки данных в уравнение (2.1), получим

$$\frac{\partial^2 Q_A^*}{\partial x^2} + \frac{\partial^2 Q_A^*}{\partial y^2} = -A_{nm} \pi^2 \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{b} \left( \frac{n^2}{a^2} + \frac{m^2}{b^2} \right), \quad (2.3)$$

где [7]

$$A_{nm} = \frac{4}{ab} \int_0^a \int_0^b Q^*(x, y) \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{b} dx dy, \quad 0 < x < a, \quad 0 < y < b.$$

Значение коэффициента  $A_{nm}$  для равномерного заданного распределения  $Q^*(x, y) = Q^* = \text{const}$ ,  $0 < x < a, 0 < y < b$ , запишется следующим образом:

$$A_{nm} = \frac{4Q^*}{ab} \int_0^a \sin \frac{n\pi x}{a} dx \int_0^b \sin \frac{m\pi y}{b} dy. \quad (2.4)$$

Заметим, что требования 1 и 2 (см. § 1), предъявляемые к аппроксимирующей функции, выполнены.

Вычислив значение коэффициента (2.4), получим

$$A_{nm} = \begin{cases} \frac{4Q^* 2a 2b}{b n\pi t\pi} = \frac{16}{n\pi t\pi^2} Q^*, & \text{если } n \text{ и } m \text{ нечетные,} \\ 0, & \text{если } n \text{ или } m \text{ четные.} \end{cases}$$

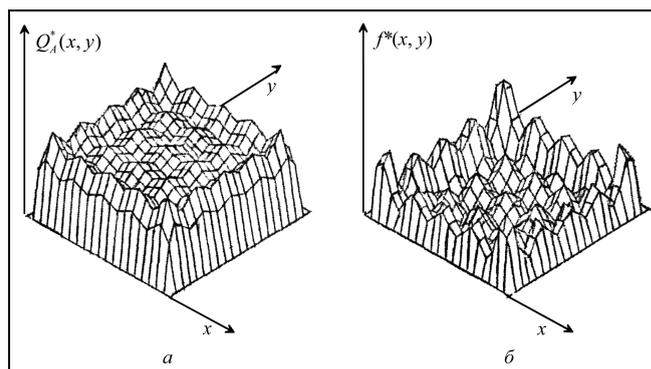


Рис. 1. Вид аппроксимирующей функции (а)  $Q_A^*(x, y)$  и соответствующего ей распределенного воздействия (б) для  $N = M = 10$

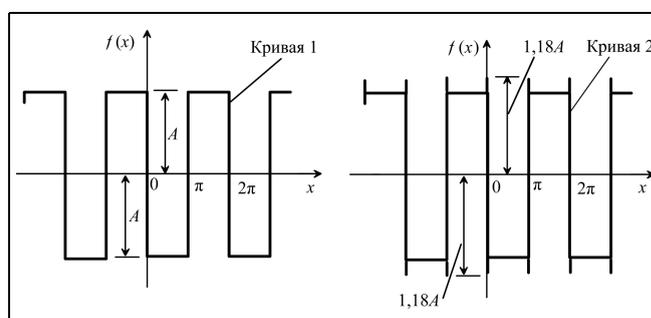


Рис. 2. Предельный «геометрический образ» полинома Фурье для одномерного случая

Для численного анализа вычисленные таким образом аппроксимирующие функции (2.2) и соответствующие им распределенные воздействия (2.1) рассчитывались с различным числом членов суммы ряда, а именно:  $N = M = 5, 10, 15, 20, 25, 30$ . Для иллюстрации результатов на рис. 1 приведены графики аппроксимирующей функции и распределенного воздействия с числом членов суммы ряда  $N = M = 10$ .

Видно, что график аппроксимирующей функции имеет несколько экстремальных значений, причем амплитуда колебаний не имеет тенденции уменьшаться до нуля при увеличении числа членов суммы ряда.

Предельным «геометрическим образом» полинома Фурье при  $n \rightarrow \infty$  в одномерном случае служит не исходная «прямоугольная» кривая 1 (рис. 2), а кривая 2 с удлиненными примерно на 18 % вертикальными отрезками, в которых слились горбы и впадины колебаний графика при  $n \rightarrow \infty$ . Это известное явление Гиббса [8].

Построим аналогичные графики для двумерной аппроксимирующей функции (2.2) при различном числе членов суммы. Поскольку графики  $Q_A^*(x, y)$

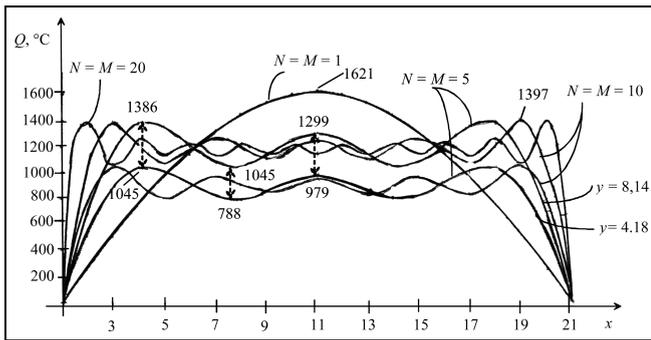


Рис. 3. Сравнительные графики температуры при  $y = \text{const}$  для  $N = M = 5; 10; 20$

в рассматриваемом нами случае симметричны, см. рис. 1, (расчет производился для пластины с параметрами  $a = b$ ), то интересующие нас графики можно строить по сечениям, т. е. по одной из координат, зафиксировав другую. На рис. 3 показаны кривые  $Q_A^*(x, y)$  для  $N = M = 5, N = M = 10$  и  $N = M = 20$  при  $y = \text{const}$ . Значения констант выбраны для самого неблагоприятного случая, когда «перерегулирование» самое большое по абсолютной величине.

Расчеты показывают, что в отличие от точек непрерывности, в которых аппроксимирующий полином Фурье при  $n \rightarrow \infty$  и  $m \rightarrow \infty$  стремится к исходной функции  $Q_A^*(x, y)$ , в точках разрыва и их окрестностях аппроксимирующий полином Фурье дает заметную погрешность даже при большом числе суммируемых членов, хотя чем больше число суммируемых членов, тем уже область вблизи точек разрыва, в которой погрешность может быть существенна.

### 3. ПРИМЕНЕНИЕ ПОЛИНОМОВ РАВНОМЕРНОГО ПРИБЛИЖЕНИЯ ДЛЯ РАСЧЕТА АППРОКСИМИРУЮЩЕЙ ФУНКЦИИ

Результаты суммирования тригонометрических рядов разрывных функций необходимо критически оценивать, учитывая явление Гиббса, так как дефект сходимости имеет место даже при числе членов ряда, стремящемся к бесконечности. Кроме того, при использовании метода подстановки в двумерном случае необходимо брать вторую производную от аппроксимирующей функции. Если функция  $Q_A^*(x, y)$  — многоэкстремальная, то, как в одномерном, так и в двумерном случае характер колебаний сохраняется. Но при увеличении размерности задачи (двумерный случай) разности

между максимальными и минимальными значениями амплитуд возрастают. Более того, эти разности будут возрастать и при увеличении числа членов ряда. Это может привести к тому, что распределенное управление  $f^*(x, y)$ , полученное методом подстановки, будет не пригодно для практического использования.

Важно аппроксимировать функцию так, чтобы амплитуда этих колебаний не превышала некоторой заданной величины. Существуют аппроксимирующие полиномы, удовлетворяющие такому условию. Они называются полиномами равномерного приближения, или (если требуется, чтобы все максимальные отклонения были одинаковыми) равномерным чебышевским приближением (или просто чебышевским приближением). Ценно, что при аппроксимации ими разрывных функций исключается или ослабляется явление Гиббса. Кроме того, в ряде случаев удается удовлетворительно аппроксимировать функцию полиномом меньшего порядка, чем при аппроксимации ее полиномом Фурье (при той же точности) [8].

Общее свойство всякого полинома равномерного приближения состоит в том, что с увеличением порядка  $N$  качество аппроксимации улучшается. Путем умножения каждого из коэффициентов Фурье  $C_n$  на множители сходимости  $\rho_n(N)$  достигается превращение полинома Фурье в полином равномерного приближения.

Совершенно ясно, что различные множители сходимости дают количественно равномерное неодинаковое приближение, и выбор метода зависит от требований к точности в конкретной задаче.

Применение множителей сходимости дает хорошие результаты для функций, зависящих от одной переменной [8]. Для двумерного случая нами проведены аналогичные расчеты аппроксимирующей функции  $Q_A^*(x, y)$  в виде полинома Фурье с применением множителей сходимости. Были выбраны множители Бернштейна—Рогозинского:

$$\rho_{B, n}(r, N) = \left( \cos \frac{n\pi}{2N+1} \right)^r.$$

Анализируя результаты расчетов для одномерного случая, приведенные в работе [8, с. 363], можно видеть, что применение множителей Бернштейна—Рогозинского дает наименьшие колебания и, кроме того, наилучшим образом аппроксимирует разрывную функцию (в смысле выделения крутизны фронтов) при том же числе членов в полиноме.

Аппроксимирующая функция (2.2) с применением множителей Бернштейна—Рогозинского  $\rho_{B, n}$



и  $\rho_{B,m}$  для рассматриваемого двумерного случая запишется следующим образом:

$$Q_A^*(x, y) = \sum_{n=1}^N \sum_{m=1}^M A_{nm} \left( \cos \frac{n\pi}{2N+1} \right)^r \left( \cos \frac{m\pi}{2M+1} \right)^r \times \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{b}.$$

Для оценки эффекта применения множителей сходимости, расчеты проводились для тех же самых значений числа членов суммы ряда, что и в предыдущем случае (без применения множителей сходимости):  $N = M = 5, 10, 15, 20, 25, 30$ . На рис. 4 показан вид аппроксимирующей функции  $Q_A^*(x, y)$  и распределенного воздействия  $f^*(x, y)$ , вычисленных с применением множителей сходимости Бернштейна – Рогозинского с показателем степени  $r = 2$  и числом членов суммы  $N = M = 10$  (сравните с рис. 1).

На рис. 5 представлена зависимость  $Q_A^*(x, y)$  при фиксированном  $y$ . Требуемая температура распределения  $Q_A^*(x, y) = 1000^\circ\text{C}$ , число членов суммы ряда  $N = M = 20$ . Построена трубка по сечению  $y = 3$  с минимальной температурой и по сечению  $y = 2$  с максимальной температурой. Все другие значения температуры распределения расположены внутри этой трубки (для примера показано сечение при  $y = 11$ ).

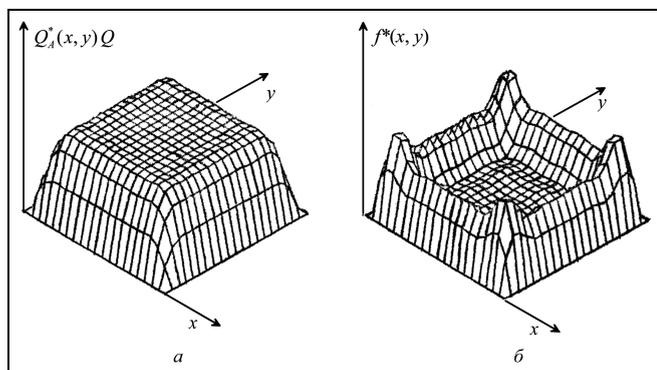


Рис. 4. Вид аппроксимирующей функции (а) и соответствующего ей распределенного воздействия (б) для  $N = M = 10$ , рассчитанные с применением множителей сходимости

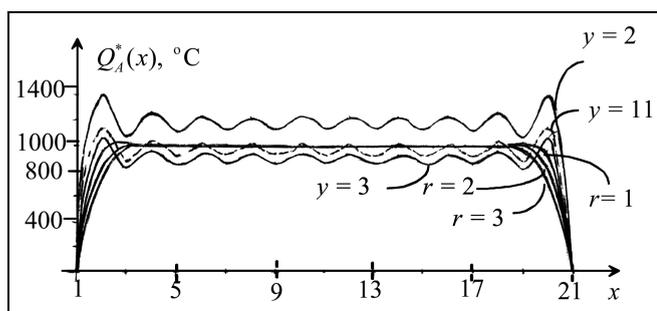


Рис. 5. Графики температуры при фиксированном  $y$

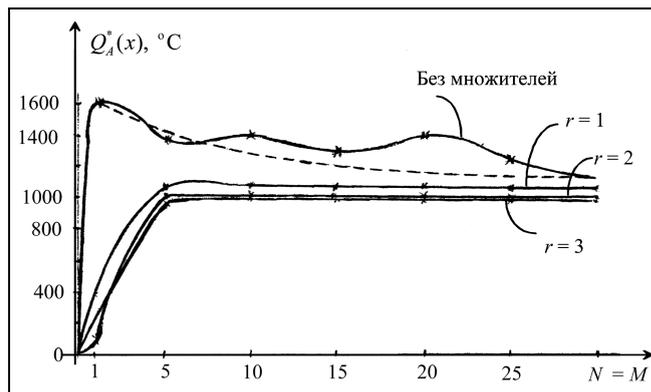


Рис. 6. Зависимость расчетной температуры от числа членов суммы ряда

чению  $y = 2$  с максимальной температурой. Все другие значения температуры распределения расположены внутри этой трубки (для примера показано сечение при  $y = 11$ ).

На рис. 6 представлены зависимости значений температуры в самой «тяжелой» точке распределения от числа членов суммы ряда.

Кривые, полученные при расчете без множителей сходимости, подтверждают выводы, сделанные для одномерного случая [8]: наличие значительные колебания аппроксимирующей функции. Максимальная амплитуда отклонения от требуемого уровня температуры (в точке с координатами  $x = 2, y = 2$ ) составила  $A_{\max} = 39,1 \%$ . Применение множителей сходимости дало результирующее распределение температуры практически равномерное при всех показателях степени  $r = 1, 2, 3$ , при которых были проведены расчеты. Максимальное значение  $A_{\max} = 0,6 \%$  при  $r = 1$ ;  $A_{\max} = 2,5 \%$  при  $r = 2$ , а при  $r = 3$  колебаний температуры не наблюдается. Самый же крутой фронт наблюдается при  $r = 1$ .

Анализ результатов расчета показал, что достаточно небольшого числа членов суммы ряда (например, 10, при показателе степени  $r = 1$ ) для получения довольно точной аппроксимации. При требовании крутых фронтов лучше брать несколько большее число членов суммы ряда, например, 20.

#### 4. ЧИСЛЕННЫЙ ПРИМЕР ПРИМЕНЕНИЯ МЕТОДА ПОДСТАВКИ ДЛЯ СЛУЧАЯ НЕРАВНОМЕРНОГО ПОЛЯ РАСПРЕДЕЛЕНИЯ ТЕМПЕРАТУРЫ

Рассчитаем аппроксимирующую функцию с помощью полинома Фурье для случая неравномерного желаемого распределения  $Q^*(x, y)$ , изображенного на рис. 7.

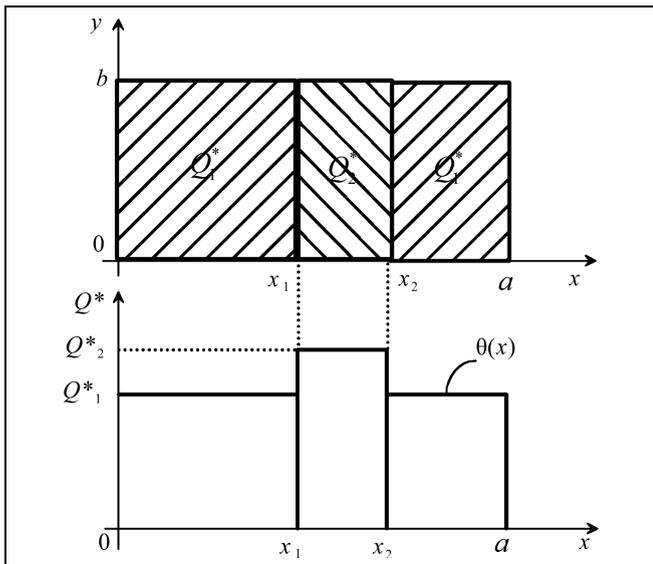


Рис. 7. Схема неравномерного распределения температуры

Аппроксимирующая функция для такой задачи записывается как  $Q_A^*(x, y) = \theta(x) \cdot 1(y)$ , где

$$\theta(x) = \begin{cases} Q_1^*, & 0 \leq x \leq x_1, \\ Q_2^*, & x_1 \leq x \leq x_2, \\ Q_1^*, & x_2 \leq x \leq a, \end{cases} \quad (4.1)$$

$$1(y) = \begin{cases} 1, & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Нетрудно видеть, что выражение (2.3) остается тем же самым, что и для равномерного распределения, а выражение (2.4) для  $A_{nm}$  запишется следующим образом

$$A_{nm} = \frac{4}{ab} \int_0^a \theta(x) \sin \frac{n\pi x}{a} dx \begin{cases} \frac{2b}{m\pi}, & \text{если } m \text{ — нечетное,} \\ 0, & \text{если } m \text{ — четное.} \end{cases} \quad (4.2)$$

В соответствии с выражением (4.1) вычислим

$$\begin{aligned} & \int_0^a \theta(x) \sin \frac{n\pi x}{a} dx = \frac{a}{n\pi} \beta, \text{ где} \\ \beta &= Q_1^* \left( \cos \frac{n\pi x_2}{a} - \cos \frac{n\pi x_1}{a} + 1 - (-1)^n \right) + \\ & + Q_2^* \left( \cos \frac{n\pi x_1}{a} - \cos \frac{n\pi x_2}{a} \right). \end{aligned} \quad (4.3)$$

В свою очередь, выражение (4.2) запишется как

$$A_{nm} = \frac{4}{ab} \frac{a}{n\pi} \beta \begin{cases} \frac{2b}{m\pi}, & \text{если } m \text{ — нечетное,} \\ 0, & \text{если } m \text{ — четное.} \end{cases}$$

или

$$A_{nm} = \begin{cases} \frac{8}{nm\pi^2} \beta, & \text{если } m \text{ — нечетное,} \\ 0, & \text{если } m \text{ — четное,} \end{cases} \quad (4.3)$$

где  $\beta$  вычисляется по формуле (4.3).

Для иллюстрации проведенного таким образом расчета на рис. 8 представлены графики функции  $Q_A^*(x, y)$ , аппроксимирующей неравномерное температурное поле и соответствующие ей графики распределенного воздействия  $f^*(x, y)$ , вычисленного по формуле (2.1), рассчитанные без множителей сходимости (рис. 8, а) и с учетом множителей Бернштейна—Рогозинского с показателями степени  $r = 1$  (рис. 8, б),  $r = 2$  (рис. 8, в),  $r = 3$  (рис. 8, г).

Расчеты (для всех представленных графиков) проводились при следующих исходных данных:  $\lambda = 104 \text{ Вт/м} \cdot \text{°C}$ ;  $\sigma = 5,7 \cdot 10^{-8} \text{ Вт/м}^2 \cdot \text{K}^4$ ;  $\varepsilon = 0,15$ ;  $Q_0 = 20 \text{ °C}$ .

## ЗАКЛЮЧЕНИЕ

Метод подстановки для расчета распределенных управлений [2] распространен на пространственно-двумерные системы.

Предложен алгоритм расчета функции, аппроксимирующей требуемое распределение температуры на поверхности объекта. В качестве примера продемонстрирована аппроксимация заданной кусочно-постоянной функции двух переменных с помощью частичной (конечной) суммы ряда Фурье для двух видов требуемого распределения.

Предложенный алгоритм реализован на специально разработанном для таких задач программном обеспечении — стационарной модели для расчета распределенных управлений.

Применение предложенного подхода позволяет получить довольно точную аппроксимацию разрывной функции (даже при небольшом числе членов суммы ряда Фурье) и достаточно крутой фронт аппроксимирующей функции.

В реальных установках, для которых разрабатывался комплекс программ для управления подвижным источником воздействия, как правило, на поверхности объекта необходимо создать не одну, а несколько областей (как минимум, три) с различ-



реходе от одной области к другой, что значительно сокращает области «сшивки» и позволяет избежать громоздких выкладок, характерных, например, для параболической аппроксимации.

## ЛИТЕРАТУРА

1. Бутковский А.Г., Пустыльников Л.М. Теория подвижного управления системами с распределенными параметрами. — М.: Наука, 1980. — 384 с.
2. Метод подстановки для расчета распределенных управлений / А.Г. Бутковский, В.А. Кубышкин, А.Г. Смирнов и др. // Автоматика и телемеханика. — 1984. — № 9. — С. 52–61.
3. Реализация распределенных управлений при помощи подвижных источников воздействия / А.Г. Бутковский, В.А. Кубышкин, Е.С. Твердохлебов, Е.П. Чубаров // Автоматика и телемеханика. — 1983. — № 4. — С. 5–12.
4. Финягина В.И. Метод подстановки в решении двумерной задачи нагрева тел с помощью подвижных источников тепла // Проблемы управления. — 2010. — № 1. — С. 57–63.
5. Применение скользящих режимов для управления объектами с распределенными параметрами с подвижным многоцикловым воздействием / А.М. Брегер, А.Г. Бутковский, В.А. Кубышкин, В.И. Уткин // Автоматика и телемеханика. — 1980. — № 3. — С. 72–83.
6. Управление распределенными системами с подвижным воздействием / Под ред. А.Г. Бутковского. — М.: Наука, 1979. — 264 с.
7. Смирнов В.И. Курс высшей математики. — М.-Л.: Гос. изд-во технико-теоретической литературы. — 1950. — Т. 2. — 622 с.
8. Заездный А.М. Гармонический синтез в радиотехнике и электросвязи. — Л.: Энергия, 1972. — 527 с.

Статья представлена к публикации членом редколлегии А.Г. Бутковским.

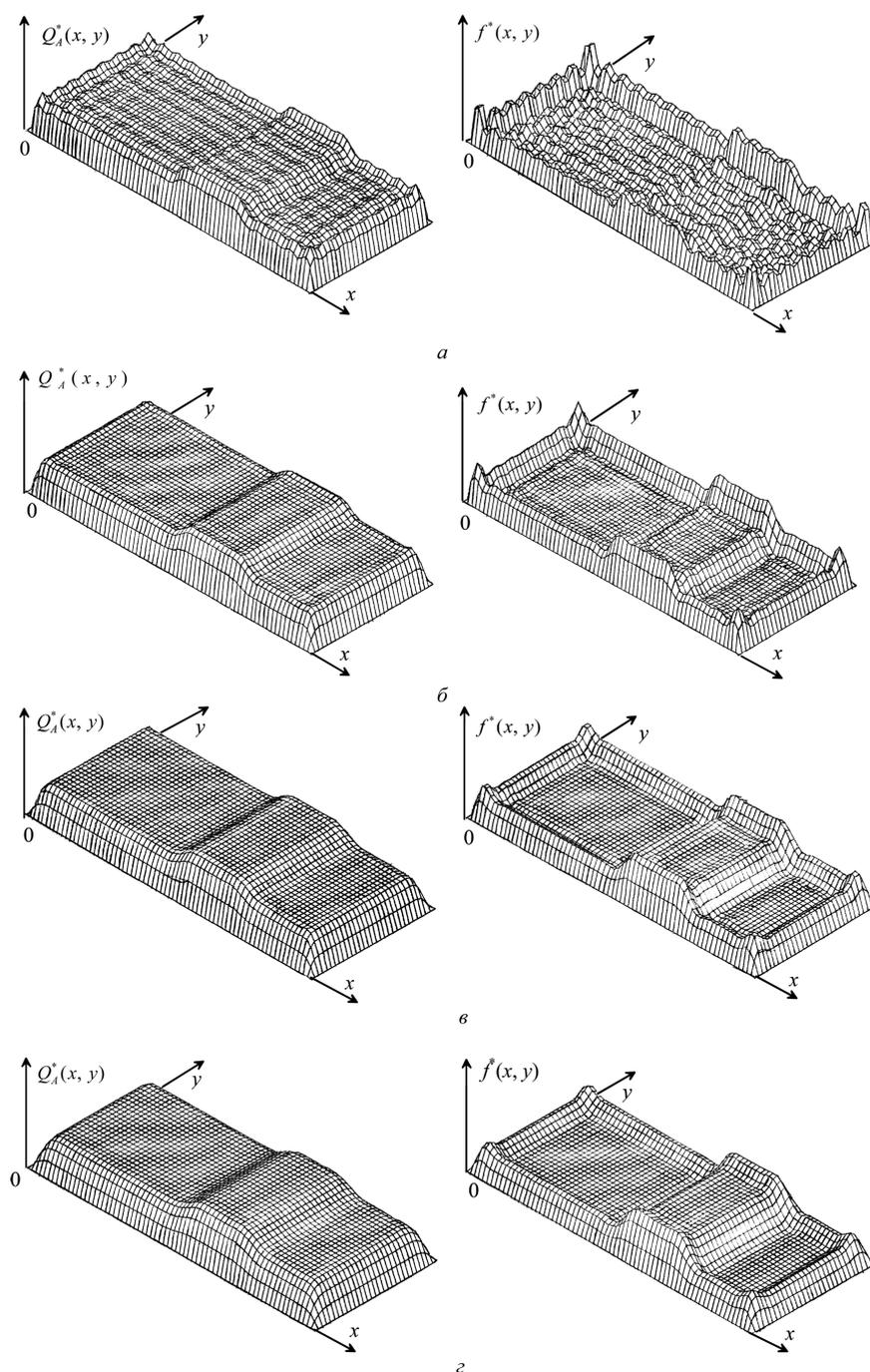


Рис. 8. Графики аппроксимирующей функции  $Q_A^*(x, y)$  и соответствующего ей распределенного воздействия  $f^*(x, y)$  для неравномерного температурного поля  $N = 35$ ,  $M = 15$ :

$a$  — расчет без множителей сходимости;  $б$  —  $r = 1$ ;  $в$  —  $r = 2$ ;  $г$  —  $r = 3$

ной температурой. В таких случаях применение предложенного способа расчета дает значительный выигрыш, поскольку можно получить довольно крутой фронт изменения температуры при пе-

Финягина Валерия Ивановна — канд. техн. наук, ст. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ☎ (495) 334-76-90, ✉ fleg@ipu.ru.

# CONTENTS & ABSTRACTS

## A REVIEW OF GAME-THEORETIC STUDIES OF INNOVATION 2 Dubina I.N.

This paper qualitatively and quantitatively characterizes the mainstream in the current innovation studies based on game theory principles and tools. The author describes the main aspects of «innovation games» analysis, including R&D and technology competition, innovation licensing, patent race, entry with a new product into a market, choice of the optimal innovation strategy (independent innovation, imitative innovation, cooperative innovation, etc.), innovation deployment, optimal sharing of innovation project outcomes, interrelation of innovation consortium participants, innovation policy and government interventions.

**Keywords:** game theory, innovation, competition, licensing, innovation consortium.

## SOLUTION OF THE GENERALIZED JOHNSON PROBLEM WITH CONSTRAINTS ON THE SCHEDULE AND TIME OF THE MACHINE. PART 2. APPROXIMATE METHODS OF SOLUTION .....12 Zak J.A.

The complex algorithm of the flow-shop problem solution using heuristics based on various hypotheses and preference rules is developed. The rules of construction of the tasks performance sequences that allow analyzing various areas of attraction of local minima are offered. The paper shows that efficiency of the constructed schedules increases due to the local variation algorithm offered. Comparison with earlier received solutions is given, the numerical example is made.

**Keywords:** optimal schedule, task performance sequence, heuristics, local variations.

## INVESTMENT AND BALANCED GROWTH IN A MODEL OF DECENTRALIZED ECONOMY .....20 Abramov A.P.

The two schemes of reaching a balanced growth are considered for the decentralized economy with investment. The first scheme is based on sales of goods. The second one is based on expected profit maximization. Both schemes are shown to solve the given task.

**Keywords:** decentralized economy, balanced growth, investment.

## MAKING INVESTMENT DECISIONS BASING UPON DISCRETE SYSTEM OPTIMAL CONTROL THEORY .....27 Pavlov O.V.

The paper presents general approach to making investment decisions basing upon discrete system optimal control theory. The formulated problem is solved with the application of maximum principle of Pontryagin. The numerical algorithm of investment optimal control is formulated. Analytical solution is obtained for the case of the fixed assets constant profitability. The criteria for making decisions of investing funds into the tangible assets are formulated.

**Keywords:** investment decisions, investment optimal control problem, discrete maximum principle of Pontryagin, numerical algorithm, criteria for making investment decisions.

## STRUCTURE-RANGING METHODOLOGY OF THE METROPOLIS HOUSING AND COMMUNAL SERVICES FUNCTIONING EFFICIENCY ESTIMATION (BY THE EXAMPLE OF MOSCOW) .....34 Dorofeyuk J.A., Mandel A.S.

The paper considers implementation of the ranging data analysis methods, the multivariant expert examination method and the expert-statistical data processing methods for efficiency estimation methodology development in the metropolis housing and communal services control problem (by the example of Moscow).

**Keywords:** ranging data analysis, multivariant expert examination, expert-statistical methods, metropolis housing and communal services, executive authority efficiency estimation system.

## MODELING OF CONFLICT SITUATION IF AGENTS HAVE DIFFERENT OPINIONS .....42 Kulivets S.G.

Two models of the game theory describing conflicts between agents in ill-structured situation were considered. Linear cognitive maps were used as a presentation of knowledge about the agents situation. Cognitive maps vary according to agents.

**Keywords:** game theory, the analysis of situations using cognitive maps, ill-structured situations, Nash equilibrium.

## THEORY FOR HIERARCHICAL MENU STRUCTURE OPTIMIZATION. ....49 Goubko M.A., Danilenko A.I.

The paper presents a formal model for hierarchical menu structure optimization. Each menu item is endowed with a probability of selection, and a discrete optimization problem is solved by choosing a menu hierarchy that minimizes the average search time. Categories in the optimal menu are proved to contain the same number of alternatives with the same relative popularity. The model covers the great variety of menu types and allows choosing the best type to meet the requirements of specific application or user category. We also suggest an automated routine for menu structure optimization that accounts for both semantic constraints and results of optimization, and illustrate it by optimizing a command menu of a mobile phone.

**Keywords:** menu-driven system, usability, depth vs. breadth, optimal hierarchy, menu design automation, menu of mobile device.

## THE INTERACTIONS OF CLIENT AND SERVER IN INTERNET-SERVICE OF REMOTE FILE PACKETS PROCESSING .....59 Asratian R.E.

Principles of the organization of network interaction in Internet service RFPS (remote file packets service), intended for the distributed control systems support are considered. The accent is made on the problem of restoration of interaction after casual or intended break of the network connection without loss of processing results. The solution of this problem by service RFPS, based on the use of remote package of data sets as a basis for creation of a stable context of the interaction which is not collapsing as a result of network failures is described.

**Keywords:** distributed control systems, Internet services, network interactions.

## OPTIMIZATION OF BACKBONE NETWORK EXPANSION ON THE BASIS OF GEOINFORMATION TECHNOLOGIES ....66 Popova O.M., Usov I.Yu.

The paper deals with an optimization problem of backbone network expansion. The structure of the software package is presented and the technique of its application is described. The indices of structural analysis of electric power systems are given. The specific features of geoinformation technology application are shown.

**Keywords:** backbone electric network, electric power system, optimization of electric network expansion, structural analysis, software package, GIS.

## DETERMINATION OF INFLUENCE COEFFICIENTS OF THE REGIME PARAMETERS ON THE ELECTRICAL POWER SYSTEM STABILITY .....74 Efimov A.E., Opalev O.L., Yadykin I.B.

The mathematical model of the electrical power system is considered in the paper. The method of the system stability analysis is based on the research of dynamics matrix of the system linearized model. The measure of stability depending on the power system parameters is introduced on the basis of the classic Lyapunov stability theory. The problem of estimating the power system stability margin is solved. The unique algorithm for computing the influence coefficients of the regime parameters on the power system stability index is presented.

**Keywords:** electrical power system, mathematical model, Lyapunov stability, margin of power system stability, dynamics matrix, stability index, influence coefficient of the regime parameter.

## CALCULATION OF APPROXIMATING FUNCTIONS OF TWO-DIMENSIONAL TEMPERATURE FIELDS BY METHOD OF SUBSTITUTION FOR CONTROL PROBLEMS OF MOBILE SOURCES OF ACTION .....79 Finyagina V.I.

Class of control problems for distributed parameter systems with mobile control is presented with solution by the method called «Method of Substitution» for calculation of distributed controls in two-dimensional plants, in particular, temperature fields. Required temperature fields are given in the form of piece-constant functions in rectangular area. Approximation of these temperature fields is realized by means of final trigonometric sums. Results of numerical simulations are presented.

**Keywords:** distributed control systems, mobile control, substitution method, 2-D temperature field, partial sum of Fourier row.