ISSN 1819-3161

ПРОБЛЕМЫ УПРАВЛЕНИЯ

5/2016



Редакционный совет

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

Редколлегия

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

Руководители региональных редакционных советов

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



CONTROL SCIENCES

Научно-технический журнал

6 номеров в год ISSN 1819-3161

Издается с 2003 года

УЧРЕДИТЕЛЬ

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

Главный редактор чл.-корр. РАН

Д.А. Новиков

Заместители главного редактора

Л.П. Боровских, Ф.Ф. Пащенко

Редактор

Т.А. Гладкова

Выпускающий редактор

Л.В. Петракова

Издатель ООО «Сенсидат-Плюс»

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

Интернет: http://pu.mtas.ru

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

Отпечатано в ООО «Адвансед солюшнз» Заказ № РВ516

Подписано в печать 3.10.2016 г.

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

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

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

Подписные индексы:

80508 и 81708 в каталоге Роспечати; 38006 в объединенном каталоге «Пресса России»

Цена свободная

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

ПРОБЛЕМЫ УПРАВЛЕНИЯ

5.2016

СОДЕРЖАНИЕ

Обзоры
Славин Б.Б. Технологии коллективного интеллекта
Математические проблемы управления
Шведов А.С. Простое доказательство робастности метода наименьших квадратов с урезанием для линейной регрессионной модели 10
Управление в социально-экономических системах
Бурков В.Н., Коргин Н.А., Новиков Д.А. Проблемы комплексирования и декомпозиции механизмов управления организационно-техническими системами
Белов М.В. Модель управления человеческим капиталом фирмы 24
Грибанова Е.Б. Решение обратных задач экономики с помощью модифицированного метода обратных вычислений
Спиро А.Г., Гольдовская М.Д., Киселева Н.Е., Покровская И.В. Сегментация и хэширование временных рядов в задачах прогнозирования на фондовом рынке
Информационные технологии в управлении
Микрин Е.А., Сомов С.К. Оптимальное оперативное резервирование информации в системах обработки данных на базе вычислительных сетей
Управление сложными технологическими процессами и производствами
Зак Ю.А. Математические модели и алгоритмы построения допустимых и оптимальных маршрутов движения и доставки грузов
Управление подвижными объектами
Потехин А.И. Логические модели объектов железнодорожной станции
* * *
Contents and abstracts



ТЕХНОЛОГИИ КОЛЛЕКТИВНОГО ИНТЕЛЛЕКТА

Б.Б. Славин

Представлен обзор литературы по технологиям коллективного интеллекта. Показано, что они существенно отличаются от краудсорсинговых технологий и могут быть рассмотрены как сорсинг эпохи знаний. Предложено описание атрибутов коллективного интеллекта, которые могут применяться для проектирования экспертных сетей и систем управления знаниями на предприятиях. Кратко описаны методы моделирования и измерения коллективного интеллекта. Показана необходимость развития теории и практики технологий коллективного интеллекта.

Ключевые слова: коллективный интеллект, краудсорсинг, коэффициент интеллектуальности, экспертные сети, социальные сети.

ВВЕДЕНИЕ

Понятие коллективного интеллекта в научной литературе стало использоваться с конца XIX в. и связано с исследованиями французских социологов того времени. В работе «О разделении общественного труда» [1] в 1893 г. Э. Дюркгейм писал о двух видах солидарности человека и общества, связанных в первом случае с общими идеями и законами поведения, а во втором — с разделением труда. К пониманию коллективного интеллекта как новой формы организации общества можно отнести и понятие ноосферы В.И. Вернадского [2]. Знаменитый фантаст Г. Уэллс в своих публицистических лекциях, изданных под названием «Всемирный мозг», в конце 1930-х гг. также говорил о важности коллективной интеллектуальной деятельности: «Думаю, что мы пока еще уделяем недостаточно внимания более тесного объединения организаций, занимающихся интеллектуальной деятельностью. Идеи объединения человечества зависят, в конечном счете, от возможности реализации единой умственной работы» [3].

Уэллс фактически предвосхитил проект Википедии, который стал возможен благодаря сетевым возможностям глобальных коммуникаций. Глобальная сеть Интернет перевела вопросы коллективного интеллекта в утилитарное русло. Ф. Хейлиген пишет: «Гораздо важнее поддерживать самоорганизацию в сети информационных потоков и таким образом содействовать повышению коллективного разума, чтобы его интеллектуальные возможности, как это положено, были больше, чем сумма сознаний пользователей Интернета» [4, *с. 92*]. Подчас, говоря о понятиях коллективного интеллекта или коллективного разума, многие исследователи подразумевают исключительно существующие глобальные сетевые проекты (см., например, обзорную статью [5]).

Известный футуролог, соучредитель и директор проекта «The Millenium Project», Дж. Гленн, заявивший в 2009 г., что коллективный интеллект станет следующей эпохой для информационных технологий, считает, что системы коллективного интеллекта (CIS) должны состоять из трех элементов: экспертов, программного и аппаратного обеспечения и из данных, информации и знания. В своей работе [6] он пишет: «Можно считать проект Wikipedia, Google, краудсорсинг, сервисы усредненных экспертных оценок, модели интеллектуального роя, а также инструменты predict markets примерами систем коллективного интеллекта, но они не являются примерами CIS по определению, предложенному в этой статье. Они производят информацию и в некоторых случаях групповые решения, но они не включают в полной мере и на систематической основе обратную связь всех трех элементов, не могут обеспечить непрерывного появления новых идей. Они не воспроизводят постоянно интеллект, только на короткое время, тогда как CIS, как и интеллект человека, постоянно возникает и меняется в процессе деятельности, приобретения опыта и влияния окружающей среды».



1. ГЛОБАЛЬНЫЕ СЕТИ И КОЛЛЕКТИВНЫЙ ИНТЕЛЛЕКТ

Безусловно, интерес к технологиям коллективного интеллекта существенно вырос с появлением коллективных веб-сервисов и таких проектов как Википедия, социальные сети. Успех Википедии заключается не только в возможности привлечения большого числа энтузиастов, но и в четкой модерации контента. Исследования, проведенные американскими учеными из Университета Карнеги — Меллон [7], показали, что эффективность создания сложного контента в Википедии возрастает с повышением компетенций редакторов, осуществляющих модерацию записей, что можно использовать для организации и других глобальных сетевых сообществ.

Японские исследователи К. Зетцу и Я. Кийоки сформулировали концепцию развития систем управления знаниями на основе применении технологий коллективного интеллекта в Интернете [8]. Они предложили рассматривать всемирную сеть как инструмент объединения интеллектуальных способностей людей через сетевые базы знаний и социальные сообщества. Информационные технологии (ИТ) и сетевые социальные сообщества можно считать своего рода инфраструктурой коллективного интеллекта. В работе японско-китайской группы ученых [9] особо отмечается «важность информационно-коммуникационных технологий в формировании коллективного интеллекта» ([9], c. 217), и авторы вводят даже понятие «суперсети знаний», состоящей из трех компонент: медиасети, сети пользователей и сети знаний, а также понятие «knoware», знания, находящегося в вычислительных ресурсах (hardware и software).

Учитывая особую роль ИТ-коммуникаций, технологии коллективного интеллекта можно определить как сетевые информационные системы, позволяющие людям эффективно и без ограничений на расстояние организовывать совместную интеллектуальную деятельность. Надо отметить, что технологии коллективного интеллекта являются одним из важных инструментов управления неявными знаниями, которые еще требуют исследований и внедрений в практику. В 2006 г. при Массачусетском технологическим институте был создан Центр исследований коллективного интеллекта (MIT Center for Collective Intelligence — CCI), B задачу которого входит исследование технологий коллективного разума и возможностей использования их на практике в различных областях бизнеса и общества. По мнению организаторов Центра «коллективный интеллект» может быть использован и как инструмент повышения организационной эффективности и производительности фирм, и как технология работы в команде,



Рис. 1. Элементы строительных блоков или «генов» коллективного интеллекта [10]

и как методология в менеджменте. Исследование коллективного разума предполагает изучение вопросов создания глобальных вычислительных знаниевых ресурсов для организации сверхчеловеческого интеллекта, изучение возможностей человеческого мозга и новых способов организации коллективной деятельности, изучение возможностей комбинирования искусственного интеллекта и интеллекта человека.

На основе исследования большого числа практических примеров использования технологий коллективного интеллекта в 2009 г. Т. Малоун (директор МІТ ССІ) с коллегами попытались выявить «геном коллективного разума» [10, 11]. В своей работе «Укрощение толпы: выявление генома коллективного интеллекта» они дали достаточно широкое определение коллективного интеллекта как организации интеллектуальной деятельности групп людей [10, c. 2] и предложили классифицировать элементы, из которых складывается геном по четырем категориям: Кто выполняет задачу? Почему он это делают? Что свершается? Как делается? (рис. 1).

Два «гена» отвечают на вопрос «Кто?»: «Иерархия» и «Сообщество» (crowd). Ген «Иерархия» подразумевает наличие организационной структуры в принятии решений при коллективной работе. Например, несмотря на демократичный принцип разработки свободного программного обеспечения, решения о включении новых модулей в релиз Линукса принимаются Л. Торвальдсом и его помощниками («лейтенантами»). Ген «Сообщество», напротив, предполагает свободное участие людей в реализации проекта, использующего коллективные технологии, свободный «вход» и «выход». «Гены», отвечающие на вопрос «Почему?», — это: «Деньги», «Любовь» и «Слава», обозначающие основные мотивы людей, участвовать в работе сетевых сообществ. На вопрос «Что?» отвечают «гены» «Создаем» и «Решаем», и на вопрос «Как?»: «гены» с названием «Индивидуально» и «Коллективно».



Таблица 1

Анализ «генома» проекта создания Линукса [10]

Пример	Что?	Кто?	Почему?	Как?	
Линукс	Создаем	Новые программные модули	Сообщество	Деньги, Любовь, Слава	Коллегиально
	Решаем	Какой из модулей включать в релиз?	Л. Торвальдс и «лейтенанты»	Любовь, Слава	Индивидуально (по иерархии)

В табл. 1 представлен пример «генома» проекта создания Линукса, собранного из перечисленных выше «генов». По аналогии с проектом Линукса авторы [10] строят «геномы» других сетевых проектов (Википедия, InnoCentive, Threadless и др.), анализируя их особенности, разнообразные схемы принятия решений (консенсус, метод средней точки и др.) и типы коллегиальной работы (собрание, конкурс, сотрудничество и т. п.). Однако, несмотря на оригинальный прием анализа глобальных сетевых проектов, основанный на аналогии с геномом человека, предложенной классификацией «генов» трудно воспользоваться в качестве инструмента повышения эффективности или организации новых проектов, использующих технологии коллективного интеллекта. Немецкий ученый Дж. Лейместер, комментируя работу Малоуна и коллег, также пишет о необходимости изучения практического инструментария и разработки приложений технологий коллективного интеллекта с применением подходов инженерии бизнеса и информационных систем (BISE — Business and Information Systems Engineering) [12, с. 247], и такие работы уже появляются (см., например, [13]).

2. КРАУДСОРСИНГ ИЛИ ТЕХНОЛОГИИ КОЛЛЕКТИВНОГО ИНТЕЛЛЕКТА

В научной литературе можно встретить точку зрения, согласно которой краудсорсинг (и его разновидности в виде технологий открытых инноваций и управления идеями) и технологии коллективного интеллекта мало различаются [14, 15]. Так, немецкие ученые в своей статье, посвященной коллективному интеллекту, заявляют, что «мы будем использовать термины «коллективный разум» и «краудсорсинг» взаимозаменяемо и в смысле «использования большой группы людей для решения конкретной проблемы или сбора полезных идей»« [14, c. 681]. Однако краудсорсинг является технологией информационного общества, инструментом самообслуживания людей, и было бы неправильно его приравнивать к системам управления знаниями и коллективным интеллектом.

Т. Грубер так писал о краудсорсинговых проектах: «я бы назвал нынешнее состояние социальной сети как собрание интеллекта. Я думаю, что преждевременно применять термин «коллективный интеллект» к таким системам, потому что они не формируют новых уровней мышления» [16, c. 4]. Говоря о технологиях коллективного интеллекта, необходимо в первую очередь рассматривать возможности увеличения интеллектуальных способностей групповой работы и только затем возможности информационной поддержки благодаря массовости пользователей. В этой связи более адекватным представляется определение систем, использующих технологии коллективного интеллекта как «систем, которые объединяют в группы необходимое число людей, имеющих собственные индивидуальные цели, но организованных таким образом, что общий интеллект и эффективность группы возрастает» [17, *с. 219*].

Особое значение в использовании технологий коллективного интеллекта имеет их соотношение с технологиями бизнес-анализа (business intelligence) и искусственного интеллекта. Как и в случае с краудсорсингом, некоторые исследователи также не делают различия между этими «интеллектуальными» технологиями. Так, автор книги «Коллективный интеллект в действии» [18] С. Алаг пишет о коллективном интеллекте как о технологии повышения интеллектуальных возможностей человека, включая сюда и коллективные технологии, и технологии интеллектуального анализа данных (Data Mining), и машинное обучение, и т. д., все, что имеет в название слово интеллект. Впрочем, часто, особенно в публицистических выступлениях, встречается и противоположная точка зрения, что технологии искусственного интеллекта не дополняют, а заменяют интеллектуальные способности человека, создавая предпосылки для создания фантастического мира машин. Недостатки обеих точек зрения заключаются в различной трактовке слова «интеллект» (intelligence). Если же под интеллектом иметь в виду не инструментарий, а мышление, причем в его социальном значении, тогда технологии искусственного интеллекта следует понимать как технологии, помогающие че-



ловеку мыслить, избавляя его от рутинной логической деятельности, и как технологии, которые позволяют складывать интеллектуальные способности людей в единый, коллективный интеллект. Т. Грубер образно описал это сочетание технологий как «встречу социальной сети с семантической» [16].

Ограниченность краудсорсинговых технологий, которые предполагают необязательность участия, можно продемонстрировать на примере работы экспертной группы. Исследуемая экспертная группа состояла из 45 экспертов, которые готовили отчет о научно-исследовательской работе, и пользовалась в качестве инструмента коллаборации сетевым мессенджером «Telegram». На рис. 2 показан объем информации (в килобайтах), сгенерированный членами экспертной группы, обозначенные инициалами по оси ординат. Нетрудно видеть, что из 45 экспертов только половина экспертов участвовала в обсуждении, остальные довольствовались ролью слушателей. Более того, свыше 80 % всей информации было сгенерировано восемью экспертами, представляющими собой менее 20 % всего коллектива. Можно сказать, что в сообществах, построенных на технологиях краудсорсинга, работает правило Парето, предполагающее, что 80 % результата достигается благодаря использованию 20 % интеллектуальных ресурсов.

Но краудсорсинговые технологии приводят не только к неравномерности загрузки экспертов, но и к неравномерности в коммуникациях. На рис. 3 показано распределение объема информации, генерируемой экспертами в сетевом мессенджере в процессе упомянутой выше групповой работы. Нетрудно видеть, что обмен информацией между экспертами носит ярко выраженный неравномерный характер и скорее напоминает быстротечные обмены мнениями, инициируемые одним из экспертов. Понятно, что такое стратифицированное общение не соответствует реальной равномерной работе экспертов, и, следовательно, сетевые мессенджеры не могут быть эффективным инструментом организации коллективной интеллектуальной деятельности. Краудсорсинговые технологии позволяют объединить большое число людей и собрать наиболее адекватную информацию, но в части организации интеллектуальной деятельности они проигрывают даже обычным корпоративным информационным системам, которые добиваются эффективности в совместной работы путем использования бизнес-процессов.

Говоря о технологиях коллективного интеллекта, необходимо в первую очередь рассматривать возможности увеличения интеллектуальных способностей групповой работы, и только затем возможности информационной поддержки благода-



Рис. 2. Распределение объема информации по участникам экспертной группы в процессе интеллектуальной деятельности



Рис. 3. Распределение сгенерированного объема информации участниками экспертной группы по времени

ря массовости пользователей. В этой связи более адекватным представляется определение систем, использующих технологии коллективного интеллекта, как «систем, которые объединяют в группы необходимое число людей, имеющих собственные индивидуальные цели, но организованных таким образом, что общий интеллект и эффективность группы возрастают» [17], с. 219].

3. ТЕХНОЛОГИИ КОЛЛЕКТИВНОГО ИНТЕЛЛЕКТА КАК ФОРМА СОРСИНГА ЭПОХИ ЗНАНИЙ

Чтобы разобраться, чем коллективный интеллект отличается от краудсорсинга и других технологий, удобно воспользоваться классификацией технологических эпох [19] (индустриальной, постиндустриальной, информационной и эпохи знаний), каждой из которых соответствует своя форма сорсинга (инсорсинг, аутсорсинг, краудсорсинг и коллективный интеллект), и ответить на те же вопросы, что задавались Малоуном и коллегами

į



при построении генома коллективного интеллекта: «Кто?», «Что?», «Как?» и «Почему?» использует данную технологию (табл. 2). В случае краудсорсинга основным участником становятся обычные люди или клиенты, а предметом использования информация, которой люди либо делятся с другими, либо получают в результате данного сервиса. В технологии коллективного интеллекта участниками являются профессионалы, эксперты, которые либо создают новое знание, либо используют его в своей деятельности. Уже из этих отличительных особенностей видно, что проект Википедии является краудсорсинговым проектом, а не коллективным интеллектом. И наоборот, проект разработки Линукса является проектом коллективного интеллекта, а не краудсорсинговым проектом.

Основным инструментом работы в краудсорсинговых проектах является единый сетевой информационный ресурс и индивидуальная работа с ним через «личный кабинет». В случае использования технологии коллективного интеллекта сетевой ресурс служит лишь вспомогательной коммуникационной площадкой для организации коллективной работы, который легко может быть заменен обычной электронной почтой. Наиболее четко различие между технологиями проявляется в мотивации пользователей. В табл. 2 указаны лишь основные мотивации, которые, конечно же, дополняются другими. В случае инсорсинга основной мотивацией является зарплата сотрудников (хотя и дополнительные мотивации тоже имеются). При аутсорсинге основная мотивация состоит в прибыли аутсорсинговой компании, что делает эту технологию более эффективной в условиях тиражируемости услуги, чем инсорсинг. В случае краудсорсинга мотивацией пользователей служит либо их потребность в доступе к информации (например, в случае банковского электронного самообслуживания), либо социальный успех. Безусловно, и владение информации, и социальный успех

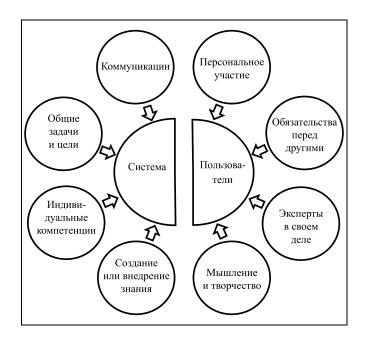


Рис. 4. Атрибуты коллективного интеллекта

могут быть и будут монетизированы пользователем, но уже опосредованно, через работу с информацией. Основная же мотивация участника проектов коллективного интеллекта заключается в потребности в коллективной знаниевой деятельности, а оплата, признание и прочие виды мотивации (см. более подробно в работе [20]) вторичны.

Различия между моделями сорсинга позволяют выявить основные атрибуты коллективного интеллекта, представленные на рис. 4. Коллективный интеллект можно представить в виде системы коллективного интеллекта (левая часть рисунка) и пользователей этой системы. Система коллективного интеллекта должна включать в себя, как и краудсорсинговые платформы, сетевой инструмент коммуникаций (для организации коллектив-

Таблица 2

Геном сорсинга

	Инсорсинг	Аутсорсинг	Краудсорсинг, самообслуживание	Коллективный интеллект
Кто? Персонал	Сотрудники предприятия	Сотрудники фирм-партнеров	Клиенты, население	Эксперты
Что? Цель	Выполняют часть работы	Оказывают услуги	Получают или дают информацию	Создают и используют знание
Как? Процесс	В рамках биз- нес-процесса	В рамках рыночных договоренностей	Интернет-самообслуживание	Коллективная работа
Почему? Мотивация	Зарплата	Прибыль	Необходимость в информации, слава	Удовлетворение в творчестве комплексных механизмов

ĥ



ной работы пользователей); иметь общие для всех задачи и цели (не все краудсорсинговые проекты имеют общие цели); определять индивидуальные компетенции (это отличает коллективный интеллект от краудсорсинговых проектов); поддерживать появление или внедрение знания. Последний атрибут систем коллективного интеллекта также отличает их от краудсорсинга. В частности, технология модерации проекта Википедии не только не поддерживает появление нового знания, а наоборот — все, что не подтверждается более старыми и авторитетными источниками, запрещает к публикации. Википедия служит инструментом хранения и передачи знаний, но не их производства или использования.

Существенные различия в атрибутике имеются и со стороны пользователей систем коллективного интеллекта. В отличие от краудсорсинговых проектов, где возможно даже анонимное участие, в системах коллективного интеллекта участие персонифицировано. Более того, пользователь систем коллективного интеллекта должен брать на себя определенные обязательства (так в экспертных сетях пользователь подписывает соглашение, препятствующее распространению инсайдерской информации [21]) в отличие от «свободного входа» и «свободного выхода» в краудсорсинговых проектах [22]. Но при этом пользователи систем коллективного интеллекта не являются ни наемными работниками, ни бизнесменами, как в случае инсорсинга или аутсорсинга. Пользователи систем коллективного интеллекта — это избранные люди, профессионалы в своем деле, прошедшие определенный отбор для участия в отличие от демократичных процедур в краудсорсинговых проектах. Еще одно отличие пользователей систем коллективного интеллекта — это их деятельность, заключающаяся исключительно в творческой и интеллектуальной работе.

4. МЕТОДЫ МОДЕЛИРОВАНИЯ И ИЗМЕРЕНИЯ КОЛЛЕКТИВНОГО ИНТЕЛЛЕКТА

Несмотря на то, что понятие коллективного интеллекта используется очень широко, работ по построению математических моделей и измерению коллективного интеллекта очень немного. В качестве одних из первых работ в области расчета коллективного коэффициента интеллектуальности (Collective Intelligence Quotient — CIQ) стоит упомянуть работы польского ученого Т. Цзубы (см. например, [23]), который использовал квази-хаотическую компьютерную модель, применяемую в биологии и социологии. Подобно тому, как броуновская частица движется благодаря хаотическим столкновениям с ней молекул, так и задачи, которые не могут быть решены одним челове-

ком, будут решаться группой людей даже в условиях их случайных коллабораций: «Мы говорим, что коллективный интеллект возникает в результате взаимодействия и сосуществования, если есть хотя бы одна проблема, которая может быть решена индивидом только при поддержке группы, либо совместно» [23, с. 492]. Статистическое моделирование решения групповых задач позволяет показывать возможности коллективного интеллекта, однако коэффициент интеллекта, рассчитанный таким образом, никак не связан с традиционным расчетом IQ (Intelligence Quotient) для индивида. Цзуба сам заявляет: «Парадоксально, но оценка коллективного интеллекта социальных структур получается проще оценки IQ одного существа» [23, с. 490]. Однако применяемый Цзубой подход интересен лишь с иллюстративной точки зрения, поскольку никак не опирается на моделирование интеллекта индивида. Фактически, предложенная модель описывает любой групповой процесс, связанный или не связанный с интеллектом.

Математическое описание коллективных процессов, которые могут использоваться для моделирования интеллектуальной деятельности, можно найти в книге Д. Вулперта «Теория коллективного интеллекта» [24]. Автор вводит функции полезности для членов коллектива и всего коллектива в целом и ищет условия, при которых частный интерес не противоречит, а наоборот способствует реализации интереса коллективного. По аналогии с методами теории игр, Вулперт вводит функции, переменные которых вполне абстрактны, и могут описывать как связи, параметры, так и самих участников коллективной деятельности. Аналогичные методы, названные им COIN (COllective INtelligence), Вулперт с коллегами применял и ранее для описания различных схем коллективного взаимодействия (см., например, работу [25]).

Голландский ученый М. Шют предпринял попытку [26] описать общий подход к моделированию коллективного интеллекта (CI). Он предложил учитывать в моделях следующие свойства CI: мультиагентность, свойства пчелиного роя, сложность, адаптивность и самоорганизуемость, и использовать имитационное моделирование, позволяющее проводить численные эксперименты подобно тому, как это делается в эволюционных (Evolutionary Methods) и коэволюционных (Co-evolution) алгоритмах оптимизации популяций [27]; в обучающих (Learning Classifier Systems) алгоритмах [28]; в нейроэволюционных (Neuro-Evolution) алгоритмах, использующих технологии нейронных сетей [29]; в алгоритмах COIN, описанных выше; в алгоритмах, описывающих движение роя (Particle Swarm Optimisation) — см., например, работу [30], и др.



Интересное исследование, не связанное, правда, с построением какой-либо модели, в области измерения коллективного IQ пользователей краудсорсинговой платформы было проведено группой ученых из Кембриджа [31]. Исследователи поставили перед собой задачу определения IQ сообществ пользователей сетевой коллаборационной платформы «Amazon's Mechanical Turk», которые решали задачи за вознаграждение. Для решения были приготовлены тестовые наборы, близкие по форме к стандартным тестам IQ, что позволило сравнить изменение коллективного IQ как при работе одного, так и при работе нескольких участников. В данной работе, благодаря возможностям краудсорсинговой платформы, исследовалась зависимость эффективности решения тестов не только от числа участников, но и от схемы вознаграждения за решение задач и от репутации участников.

Было показано, что коллективный IQ (называемый в статье, правда, Crowd IQ) растет с ростом числа участников, при этом отбор пользователей, имеющих более высокую репутацию, повышает число решений тестовых наборов в единицу времени. Интересно, что зависимость от мотивации оказалась немонотонной, она практически не менялась при низкой и высокой оплате за работу. Однако использование коммерческой краудсорсинговой платформы с реальными пользователями для проведения тестов, несмотря на удобство и дополнительные возможности, имело и недостаток, поскольку не позволяла авторам более детально изучить зависимость коллективного IQ от состава участников. Несмотря на это, полученные результаты, безусловно, интересны для понимания возможностей коллективного интеллекта, пусть даже и неорганизованного сообщества.

Важные результаты в области моделирования и изучения технологий коллективного интеллекта получены в работе [32]. Для описания математической модели расчета коллективного IQ авторы вводят понятие актора, участника тестов, который в системах коллективного интеллекта выступает как в роли решателя задач, так и в роли оценщика чужих решений. Развиваемый метод эволюционного согласования подразумевает такую организацию работ акторов, которая приводит к существенному увеличению эффективности работы группы благодаря определению более компетентных участников и перераспределению роли в групповой работе. По сути дела, эволюционное согласование, применяемое в качестве методики решения тестов, позволяет в процессе тестирования интеллектуальной деятельности выявлять компетенции членов группы, чтобы более эффективно использовать индивидуальный интеллект для решения групповых задач.

Измерять коллективный IQ можно аналогично измерению индивидуального IQ (см. например, работу [33]), только с учетом того, что тесты будут решаться группой. Важное условие групповой работы решения тестов заключается в возможности комбинировать (распределять) задачи между участниками.

ЗАКЛЮЧЕНИЕ

Изучение коллективного интеллекта, как уже говорилось, началось с исследования психологии групп и толпы. Но именно психологическая составляющая коллективного интеллекта пока еще меньше всего исследована. Чаще всего изучение психологии групп касается изучения роли лидеров, различных вариантов консенсуса в решении тех или иных задач [34, 35], исследований же коллектива как целого очень мало. В этой связи интересна работа А. Вулли с коллегами (при участии Т. Малоуна) по исследованию корреляции эффективности групповой работы и интеллектуальных способностей испытуемых [36, 37]. Было показано, что эффективность групповой работы возрастает с увеличением компетентности членов группы, но при этом наличие ярких лидеров отнюдь не увеличивает, а даже снижает производительность групповой работы по сравнению с группой более-менее равных по компетенциям испытуемых. Авторы предложили рассматривать коллективный интеллект состоящим из двух составляющих: «снизу — вверх», которая как раз описывает влияние членов группы на эффективность, и «сверху — вниз», описывающая влияние коллектива в целом на работу его членов. Именно вторая составляющая, включающая в себя атмосферу сетевой коллективной работы и возможности сетевого брейнсторминга [38], до сих пор наименее изучена.

Развитие социальных сетей, которые стали уже неотъемлемой частью не только повседневной жизни современного человека, но и экономики, потребовало серьезных исследований и моделирования сетевых взаимосвязей с использованием метрик отношений между пользователями сетей. Такими метриками могут быть, например, уровни доверия или влияния (см. например, книгу [39]). Вместе с тем, наряду с социальными сетями получают распространение экспертные сети и сообщества (см., например, статью [21]), доверие и влияние в которых в основном связано с компетенциями их членов. В случае сетевых экспертных сообществ необходимо уметь моделировать групповое взаимодействие экспертов, учитывающее их компетентностные возможности и максимизирующее совокупные показатели интеллекта.



ЛИТЕРАТУРА

- Дюркгейм Э. О разделении общественного труда. М.: Канон, 1996. — 430 с.
- Вернадский В.И. Биосфера и ноосфера. М.: Айриспресс, 2004. 576 с.
- 3. Wells H.G. World Brain. South Australia: University of Adelaide, 1938. URL: https://ebooks.adelaide.edu.au/w/wells/hg/world brain/ (дата обращения: 30.08.2016).
- 4. *Хейлиген* Ф. Концепция глобального мозга / В кн.: Рождение коллективного разума: О новых законах сетевого социума и сетевой экономики и об их влиянии на поведение человека. М.: Ленанд, 2014. С. 288.
- Weiss A. The Power of Collective Intelligence // Collective Intelligence. September 2005. P. 17—23.
- Glenn J. Collective intelligence systems and an application by The Millennium Project for the Egyptian Academy of Scientific Research and Technology // Technological Forecasting and Social Change. — 2015. — Vol. 97, August. — P. 7—14.
- Kittur A., Lee B., Kraut R.E. Coordination in Collective Intelligence: The Role of Team Structure and Task Interdependence / Studying Wikipedia. Boston, 2009. Vol. CHI 2009, April 4—9. P. 1490—1504.
- Zetssu K., Kiyoki Y. Towards Knowledge Management Based on Harnessing Collective Intelligence on the Web / EKAW 2006. Berlin, 2006. – Vol. LNAI 4248. – P. 350–357.
- Luo S., Xia H., Yoshida T., Wang Z. Toward collective intelligence of online communities: a primitive conceptual model // J Syst Sci Syst. — 2009. — Vol. 18, N 2. — P. 203—221.
- Malone T.W., Laubacher R., Dellarocas C. Harnessing Crowds: Mapping the Genome of Collective Intelligence, MIT Center for Collective Intelligence Massachusetts Institute of Technology. — Cambridge, MA, Working Paper No. 2009-001, 2009.
- Male S.A., Bush M.B., Chapman E.S. Perceptions of competency deficiencies in engineering graduates // Australasian Journal of Engineering Education. 2010. Vol. 16, N 1. P. 55—68.
- 12. *Leimeister J.M.* Collective Intelligence // Business & Information Systems Engineering. 2010. N 4. P. 245—248.
- 13. Gregg D.G. Designing for collective intelligence // Communications of ACM. 2010. Vol. 53. N 4. P. 134—138.
- Buecheler T., Sieg J., Fuchslin R., Pfeifer R. Alife XXII Conference // Crowdsourcing, Open Innovation and Collective Intelligence in the Scientific Method: A Research Agenda and Operational Framework. Odense, Denmark, 2010. P. 679—686.
- Bothos E., Apostolou D., Mentzas G. Collective intelligence for idea management with Internet-based information aggregation markets // Internet Research. — 2009. — Vol. 19, N 1. — P. 26—41.
- 16. Gruber T. Collective knowledge systems: Where the Social Web meets the Semantic Web // Web Semantics: Science, Services and Agents on the World Wide Web. 2008. Vol. 6. P. 4—13.
- 17. Lykourentzou I., Vergados D., Kapetanios E., Loumos V. Collective Intelligence Systems: Classification // Journal of emerging technologies in web intelligence. 2011. Vol. 3, N 3. P. 217—226.
- Alag S. Collective Intelligence in Action. Manning Publications, 2009. 424 p.
- 19. Славин Б.Б. Взаимосвязь этапов развития информационных технологий и экономики // Информационное общество. 2015. № 6. С. 4—13.
- Максимова Е.В. Мотивация экспертов к работе в профессиональной сети // В кн.: Рождение коллективного разума. — М.: ЛЕНАНД, 2013. — С. 230—244.

- 21. *Славин Б.* Современные экспертные сети // Открытые системы. 2014. № 7. С. 30—33.
- Geerts S. Discovering Crowdsourcing. Theory, Classification and Directions for use. — Eindhoven: TUE. Department Industrial Engineering and Innovation Sciences, 2009. — 117 p.
- Szuba T. A formal definition of the phenomenon of collective intelligence and its IQ measure // Future Generation Computer Systems. — 2001. — N 17. — P. 489—500.
- 24. *Wolpert D.H.* Theory of Collective Intelligence. Moffet Field, CA: NASA Ames Research Center, 2003. 64 p.
- Tumer K., Wolpert D. Collective Intelligence and Braess' Paradox. Moffett Field, CA: AAAI—00 Proceedings, 2000. 6 p.
- Schut M.C. On model design for simulation of collective intelligence // Information Sciences. 2010. N 180. P. 132—155.
- 27. Eiben A., Schut M., Toma T. Comparing communication protocols in evolving agent societies: cell phones versus shouting // Evolvability and Interaction: Evolutionary Substrates of Communication Signalling and Perception in the Dynamics of Social Complexity. 2003. Vol. TR—393. P. 22—28.
- Butz M. Rule-Based Evolutionary Online Learning Systems. Springer, 2006. — 277 p.
- Stanley K., Bryant B., Miikkulainen R. Real-time neuro-evolution in the nero video game // IEEE Transactions Evolutionary Computation. 2005. Vol. 9, N 6. P. 653—668.
- Engelbrecht A. Fundamentals of Computational Swarm Intelligence. Wiley, 2005. 672 p.
- Kosinski M., Bachrach Y., Kasneci G., Van-Gael J., Graepel T. Crowd IQ: measuring the intelligence of crowdsourcing platforms // Association for Computing Machinery. — 2012. — June 22. — P. 151—160.
- 32. Протасов В., Потапова 3., Осипчук О. Сертификация экспертов и определение относительной цены задачи в зависимости от ее сложности // Тр. XX Байкальской всерос. конф. «Информационные и математические технологии в науке и управлении». Иркутск, 2015. Т. 2. С. 164—175.
- Flynn J.R. Searching for justice: The Discovery of IQ Gains Over Time // American Psychologist. — 1999. — Vol. 54, N 1. — P. 5—20.
- Dyer J., Ioannou C., Morrell L., et al. Consensus decision making in human crowds // Animal Behaviour. 2008. Vol. 75. P. 461—470.
- 35. *Dyer J., Johansson A., Helbing D.*, et al. Leadership, consensus decision making and collective behaviour in humans // Phil. Trans. R. Soc. B. 2009.— Vol. 364. P. 781—789.
- Woolley A., Chabris C., Pentland A., et al. Evidence for a Collective Intelligence Factor in the Performance of Human Groups // Science. October 2010. N 330. P. 686—688.
- 37. Woolley A., Aggarwa I., Malone T. Collective Intelligence and Group Performance // Current Directions in Psychological Science. 2015. Vol. 24, N 6. P. 420—424.
- Raikov A.N. Holistic Discourse in the Network Cognitive Modeling // Journal of Mathematics and System Science. 2013. N 3. — P. 519—530.
- 39. *Губанов Д.А., Новиков Д.А., Чхартишвили А.Г.* Социальные сети: модели информационного влияния, управления и противоборства. М.: Изд-во физ.-мат. лит., 2010. 228 с.

Статья представлена к публикации членом редколлегии A.A. Дорофеюком.

УДК 519.233

ПРОСТОЕ ДОКАЗАТЕЛЬСТВО РОБАСТНОСТИ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ С УРЕЗАНИЕМ ДЛЯ ЛИНЕЙНОЙ РЕГРЕССИОННОЙ МОДЕЛИ

А.С. Шведов

В классической линейной регрессионной модели остатки предполагаются распределенными нормально. Но реальные данные редко в точности соответствуют предположениям классической модели. При этом даже единственное резко отличающееся наблюдение может очень сильно повлиять на оценку параметров регрессии. Одним из методов робастной регрессии с высокой пороговой точкой является метод наименьших квадратов с урезанием. Дано новое доказательство теоремы о величине пороговой точки для этого метода, значительно более простое, чем оригинальное доказательство.

Ключевые слова: робастная регрессия, метод наименьших квадратов с урезанием, пороговая точка.

ВВЕДЕНИЕ

Робастные статистические методы в значительной степени созданы авторами книг [1, 2] и играют важную роль в приложениях. Современные методы робастной регрессии представлены, например, в книгах [3—5].

К работам по робастным методам примыкают работы по методам выявления резко выделяющихся наблюдений. Робастность можно понимать как уменьшение (или снятие) влияния резко выделяющихся наблюдений. При этом природа резко выделяющихся наблюдений может быть любой, от ошибок по небрежности до представления каких-то очень важных эффектов. В дальнейшем, для определенности выражений, под резко выделяющимися наблюдениями будем понимать ошибочные данные. Хотя для существа математических результатов это замечание значения не имеет.

Одним из основных показателей робастности статистического метода является пороговая точка, которая показывает, какую часть статистических данных можно испортить сколь угодно сильно, и при этом получающийся ответ будет все равно «иметь отношение к делу».

Например, чтобы одним показателем представить положение набора точек на прямой, можно воспользоваться средним арифметическим. Но

тогда, испортив всего одну точку из этого набора, можно добиться того, чтобы ответ изменился сколь угодно сильно. Это пример статистического метода с низкой пороговой точкой. Но если в качестве показателя взять выборочную медиану, можно испортить до половины точек из набора, и ответ будет все равно не очень сильно отличаться от правильного (остается вопрос, что называть правильным ответом). Это пример робастного статистического метода, статистического метода с высокой пороговой точкой.

Метод наименьших квадратов (МНК), наиболее распространенный метод построения регрессионных зависимостей, имеет низкую пороговую точку. Параметр регрессии можно изменить сколь угодно сильно, изменив всего одну точку в наборе наблюдений.

Отметим, что с проблемой робастности тесно связана проблема наличия тяжелых хвостов у распределения вероятностей для ошибок регрессии. Те большие выбросы, которые маловероятны при нормальном распределении ошибок, становятся вероятными при распределениях с тяжелыми хвостами. По этим вопросам см., например, статью [6].

Одним из первых робастных методов построения регрессионных зависимостей является МНК с урезанием, подробно представленный в книге [3]. Состоит этот метод в следующем.



Пусть $y_1, ..., y_n$ — объясняемые переменные, действительные числа; $x_1, ..., x_n$ — регрессоры, p-мерные вектора. При каждом $\theta \in R^p$ рассматриваются ошибки

$$r_i(\theta) = y_i - \langle \theta, x_i \rangle, \quad i = 1, ..., n,$$

где через $\langle \cdot, \cdot \rangle$ обозначено скалярное произведение. Пусть эти ошибки упорядочены по возрастанию абсолютной величины

$$|r_{(1)}(\theta)| \le |r_{(2)}(\theta)| \le \dots \le |r_{(n)}(\theta)|.$$

(Порядок, разумеется, разный при разных θ .) Выбирается натуральное число h, удовлетворяющее условию

$$\left\lceil \frac{n}{2} \right\rceil < h < n$$

(через [·] обозначена целая часть действительного числа), и строится функция

$$L(\theta) = \sum_{i=1}^{h} r_{(i)}^{2}(\theta).$$

Оценка параметров регрессии, построенная методом наименьших квадратов с урезанием, имеет вид

$$\hat{\theta}_{LTS} = \underset{\theta \in R^p}{\arg \min} L(\theta).$$

Через Z обозначим набор точек (y_1, x_1) , ..., (y_n, x_n) . Пусть m < n. Через $\Pi_m(Z)$ обозначим множество таких наборов Z', состоящих из точек (y_1', x_1') , ..., (y_n', x_n') , для которых, по крайней мере, n-m точек содержатся в наборе Z.

Пусть функция T каждому набору точек Z' ставит в соответствие параметр регрессии θ . Пороговой точкой для функции T при заданном наборе Z называется величина $\varepsilon_n(T,Z)=m^*/n$, где m^* — минимальное из чисел m, обладающих тем свойством, что

$$\sup_{Z'\in\Pi_{\infty}(Z)}||T(Z')||=\infty.$$

Здесь и далее $\|\theta\| = \langle \theta, \theta \rangle^{1/2}$.

В частности, для обычного МНК $\varepsilon_n(T,Z) = 1/n$. В книге [3, *c.* 112—134] при условии, что точки $x_1, ..., x_n$ находятся в некотором «общем положе-

нии», доказывается **Теорема.** *При*

$$h = \left\lceil \frac{n}{2} \right\rceil + \left\lceil \frac{p+1}{2} \right\rceil$$

пороговая точка для МНК с урезанием

$$\varepsilon_n(T, Z) = \frac{1}{n} \left(\left[\frac{n-p}{2} \right] + 1 \right). \, \bullet$$

Это означает, что пороговая точка асимптотически равна 0,5 при $n \to \infty$ (т. е. является максимально высокой, можно испортить до половины всех точек из набора наблюдений).

Доказательство этой теоремы распадается на две части [3]. Отдельно доказывается, что $\frac{1}{n}\left(\left[\frac{n-p}{2}\right]+1\right)$ является оценкой снизу для пороговой точки, и что эта же величина является оценкой сверху для пороговой точки. Утверждение, что данная величина представляет собой оценку сверху для пороговой точки, доказывается достаточно коротко [3, c. 125] и, скорее всего, имеет значение лишь как красивый математический результат. Значительно более трудным в работе [3] является доказательство утверждения, что данная величина представляет собой оценку снизу для пороговой точки (именно это утверждение и означает робастность МНК с урезанием).

Отметим, что МНК с урезанием может быть применен для практического построения регрессионных зависимостей лишь в сочетании с другими алгоритмами. (Либо в прикидочных расчетах, когда нет высоких требований к точности.) Это связано с тем, что если набор наблюдений $(y_1, x_1), ..., (y_n, x_n)$ соответствует «чистой» нормальной случайной выборке, то точность МНК с урезанием существенно ниже, чем точность обычного МНК.

Одним из распространенных методов построения робастных оценок для параметров регрессии является метод MM -оценивания, предложенный в работе [7] (см. также книгу [4]). Построение оценки этим методом состоит из трех этапов. На первом этапе строится оценка параметров регрессии некоторым методом с высокой пороговой точкой. На втором этапе с использованием этих параметров строится M -оценка для масштаба. На третьем этапе с использованием найденного значения масштаба строится M -оценка для параметров регрессии (отсюда использование двух букв « M » в названии метода). Доказывается, что пороговая точка при MM -оценивании параметров регрессии асимптотически равна 0,5.

В настоящей работе дается значительно более простое доказательство сформулированного результата о пороговой точке для МНК с урезанием для параметров регрессии, чем доказательство из [3]; речь идет об оценке снизу для $\varepsilon_n(T,Z)$, т. е. о доказательстве робастности метода наименьших квадратов с урезанием.



1. ВСПОМОГАТЕЛЬНЫЙ РЕЗУЛЬТАТ

Во всей работе будем считать, что $n \ge p$ и что выполняется условие: любые p из векторов $x_1, ..., x_n$ образуют базис пространства R^p .

Из этого условия следует, что если для p-мерного вектора θ , $\|\theta\|=1$, и для некоторого набора векторов x_{i_1} , ..., $x_{i_{p-1}}$ выполняется

$$\langle \theta, \; x_{i_1} \rangle = 0, \; ..., \; \langle \theta, \; x_{i_{p-1}} \rangle = 0,$$

то для любого вектора x_j , не входящего в этот набор, $\langle \theta, x_i \rangle \neq 0$.

Некоторым обобщением этого утверждения является

Лемма. Существует число $\delta > 0$ такое, что если для р-мерного вектора θ , $\|\theta\| = 1$, и для некоторого набора векторов x_{i_1} , ..., $x_{i_{n-1}}$ выполняется

$$|\langle \theta, x_{i_1} \rangle| \leq \delta, ..., |\langle \theta, x_{i_{p-1}} \rangle| \leq \delta,$$

то для любого вектора x_j , не входящего в этот набор, $|\langle \theta, x_i \rangle| \ge \delta$.

Доказательство. Рассмотрим в качестве набора векторов x_{i_1} , ..., $x_{i_{p-1}}$ набор x_1 , ..., x_{p-1} . Проведя ортогонализацию Грама — Шмидта, построим ортонормированную систему e_1 , ..., e_{p-1} ,

$$e_i = \sum_{k=1}^{i} \gamma_{ik} x_k, \quad i = 1, ..., p - 1.$$

Проекция вектора θ на гиперплоскость, порожденную векторами $e_1, ..., e_{p-1},$ имеет вид

$$\langle \theta, e_1 \rangle e_1 + \dots + \langle \theta, e_{p-1} \rangle e_{p-1} = \sum_{i=1}^{p-1} \left(\sum_{k=1}^i \gamma_{ik} \langle \theta, x_k \rangle \right) e_i$$

Будем считать, что все скалярные произведения $\langle \theta, x_k \rangle$ настолько малы по абсолютной величине, что длина этой проекции не превосходит $1/\sqrt{2}$.

Выберем вектор x_j , j > p-1. Поскольку этот вектор не является линейной комбинацией векторов $x_1, ..., x_{p-1}$, можно записать

$$x_j = \sum_{i=1}^{p-1} \alpha_i x_i + \beta t,$$

где t — единичный вектор ортогональный гиперплоскости, порожденной векторами $e_1, ..., e_{p-1}; \beta \neq 0$.

Длина проекции вектора θ на гиперплоскость, порожденную векторами $e_1, ..., e_{p-1}$, не превосходит $1/\sqrt{2}$. Поэтому длина проекции вектора θ на вектор t должна быть не меньше, чем $1/\sqrt{2}$. Имеем

$$\langle \theta, t \rangle = \frac{1}{\beta} \langle \theta, x_j \rangle - \frac{1}{\beta} \sum_{i=1}^{p-1} \alpha_i \langle \theta, x_j \rangle.$$

Будем считать, что все скалярные произведения $\langle \theta, x_i \rangle$ настолько малы по абсолютной величине, что

$$\left|\frac{1}{\beta}\sum_{i=1}^{p-1}\alpha_i\langle\theta,x_i\rangle\right|\leq\frac{1}{\sqrt{2}}\,-\,\frac{1}{2}\,.$$

Тогда

$$|\langle \theta, x_i \rangle| \ge |\beta|/2.$$

Дополнительно наложим условие $\delta < |\beta|/2$.

Пока утверждение леммы доказано для одного j. Такое же рассмотрение можно провести для всех j=p,...,n и выбрать минимальное требуемое δ . Затем нужно рассмотреть все существующие наборы $x_{i_1},...,x_{i_{n-1}}$ (всего

таких наборов C_n^{p-1}) и вновь выбрать минимальное требуемое δ . Лемма доказана.

2. РЕЗУЛЬТАТ О ПОРОГОВОЙ ТОЧКЕ

Выберем числа m и h так, что $m+p \le h \le n-m$, и покажем, что тогда для МНК с урезанием

$$\sup_{Z'\in\Pi_m(Z)}||T(Z')||<\infty.$$

Для набора $Z' \in \Pi_m(Z)$ остатки обозначим

$$\rho_i(\theta) = y_i' - \langle \theta, x_i' \rangle, \quad i = 1, ..., n.$$

Положим $\theta = 0$, и пусть наименьшими по абсолютной величине h остатками являются

$$\rho_{k_1}(0), ..., \rho_{k_h}(0),$$

где $k_1, ..., k_h$ — различные числа из набора 1, ..., n. В силу условия $m \le n - h$ в наборе Z' содержатся точки $(y_{j_1}, x_{j_1}), ..., (y_{j_h}, x_{j_h})$. Тогда

$$L(0) = \sum_{i=1}^{h} \rho_{k_i}^2(0) \le \sum_{i=1}^{h} r_{j_i}^2(0) \le \sum_{i=1}^{n} r_i^2(0) = \sum_{i=1}^{n} y_i^2.$$

Положим

$$C_0 = \max_{1 \le j \le n} |y_j|$$

и выберем число C_1 такое, что

$$C_1 > \frac{1}{\delta} \left(\left(\sum_{i=1}^n y_i^2 \right)^{1/2} + C_0 \right),$$

где $\delta > 0$ определяется леммой.

Возьмем $\theta \in \mathbb{R}^p$ такое, что $\|\theta\| \geq C_1$. Пусть наименьшими по абсолютной величине h остатками являются $\rho_{k_1}(\theta)$, ..., $\rho_{k_h}(\theta)$. В силу условия



 $m + p \le h$ среди точек $(y'_{k_1}, x'_{k_1}), ..., (y'_{k_h}, x'_{k_h})$ присутствуют точки $(y_{i_1}, x_{i_1}), ..., (y_{i_p}, x_{i_p}).$

Воспользовавшись тем, что для любых действительных чисел α и β выполняется неравенство

$$|\alpha - \beta| \ge ||\alpha| - \beta||$$
,

имеем

$$L(\theta) = \sum_{i=1}^{h} \rho_{k_i}^2(\theta) \ge \sum_{l=1}^{p} r_{i_l}^2(\theta) \ge \sum_{l=1}^{p} (|\langle \theta, x_{i_l} \rangle| - |y_{i_l}|)^2.$$

Пусть $\theta_0 = \theta/|\theta|$. На основании леммы хотя бы для одного из индексов, например, для i_{p} выполняется $|\langle \theta_0, x_{i_n} \rangle| \ge \delta$. Тогда

$$L(\theta) \ge (\|\theta\| \cdot |\langle \theta_0, x_{i_p} \rangle| - |y_{i_p}|)^2 \ge (C_1 \delta - C_0)^2 > \sum_{i=1}^n y_i^2.$$

Тем самым установлено, что для любого набора $Z' \in \Pi_m(Z)$ выполняется $||T(Z')|| \leq C_1$.

Нетрудно проверить, что при

$$m = \left\lceil \frac{n-p}{2} \right\rceil, \quad h = \left\lceil \frac{n}{2} \right\rceil + \left\lceil \frac{p+1}{2} \right\rceil$$

неравенства $m + p \le h \le n - m$ имеют место. Следовательно,

$$m^* \ge \left[\frac{n-p}{2}\right] + 1.$$

ЗАКЛЮЧЕНИЕ

Результат работы заключается в новом доказательстве известной теоремы о робастности метода наименьших квадратов с урезанием для линейной регрессионной модели; под робастностью понимается высокая пороговая точка, асимптотически равная 0.5 при больших размерах выборки. Это доказательство существенно проще, чем оригинальное доказательство, приводимое в книге [3].

ЛИТЕРАТУРА

- 1. Хьюбер П.Дж. Робастность в статистике. М.: Мир,
- Робастность в статистике. Подход на основе функций влияния / Ф. Хампель и др. — М.: Мир, 1989. — 512 с.
- Rousseeuw P.J., Leroy A.M. Robust regression and outliers de-
- tection. N.-Y.: Wiley, 1987. 329 p.
 4. *Maronna R., Martin D., Yohai V.* Robust statistics: Theory and methods. - Chichester: Wiley, 2006. - 403 p.
- Andersen R. Modern methods for robust regression. Los Angeles: Sage Publications, 2008. — 107 p.
- 6. Шведов А.С. Робастная регрессия с применением t-pacпределения и EM -алгоритма // Экономический журнал ВШЭ. — 2011. — Т. 15. — С. 68—87.
- 7. Yohai V.J. High breakdown-point and high efficiency robust estimates for regression // Annals of Statistics. - 1987. -Vol. 15. — P. 642—656.

Статья представлена к публикации членом редколлегии А.Г. Кушнером.

Шведов Алексей Сергеевич — д-р физ.-мат. наук, профессор, Национальный исследовательский университет

- «Высшая школа экономики», г. Москва,
- □ ashvedov@hse.

Читайте в ближайших номерах

- ✓ Байбулатов А.А., Промыслов В.Г. Аппроксимация огибающей в приложениях «Network
- Ведешенков В.А. Подход к фрагментному диагностированию компонентов цифровых систем со структурой минимального квазиполного графа (на примере графа размера 7×7)
- ✓ Горлищев В.П., Калинин Л.А., Михальский А.И. и др. Метод коррекции электрокардиографического интервала с учетом частоты сердечных сокращений
- Губанов Д.А., Чхартишвили А.Г. Влиятельность пользователей и метапользователей социаль-
- ✓ Еналеев А.К. Согласованное управление в организационных сетевых структурах
- Кирин Ю.П., Кирьянов В.В. Робастное управление технологическими процессами производства губчатого титана
- Микрин Е.А., Сомов С.К. Оптимизация резервирования информации в распределенных системах обработки данных реального времени
- Ратнер С.В. Применение сетевого анализа среды функционирования в задачах регионального экологического менеджмента
- Стенников В.О., Пеньковский А.В., Хамисов О.В. Поиск равновесия Курно на рынке тепловой энергии в условиях конкурентного поведения источников тепла
- Талагаев Ю.В. Анализ и синтез сверхустойчивых нечетких систем Такаги Сугено



УДК 334.02:519.08

ПРОБЛЕМЫ КОМПЛЕКСИРОВАНИЯ И ДЕКОМПОЗИЦИИ МЕХАНИЗМОВ УПРАВЛЕНИЯ ОРГАНИЗАЦИОННО-ТЕХНИЧЕСКИМИ СИСТЕМАМИ¹

В.Н. Бурков, Н.А. Коргин, Д.А. Новиков

Обсуждены теоретические проблемы, возникающие при исследовании комплексов механизмов управления организационно-техническими системами. Предложены расширения классификации организационно-технических систем, возникающие при переходе от рассмотрения отдельных механизмов управления к комплексным. Подробно рассмотрены две ключевые проблемы — оценка эффективности и неманипулируемости комплексного механизма и возможность их декомпозиции для оценки на уровне отдельных элементарных механизмов.

Ключевые слова: теория управления, организационно-техническая система, системный анализ, неманипулируемость.

ВВЕДЕНИЕ

Теория игр всегда служила источником новых задач для исследования операций и дискретной математики [1, 2], будь то, например, вопрос о сложности вычисления равновесия Нэша [3, 4] или порождаемая иерархической игрой Γ_1 задача на связанный экстремум [5], и многие другие, в том числе, еще нерешенные, проблемы. Научные направления, систематически исследующие математические модели взаимодействия многих рациональных агентов в условиях асимметричной информации и нетривиального регламента их взаимодействия, — *meopuя механизмов* (mechanism design) и теория активных систем [6—11] — позволили разработать конструктивные методы анализа и синтеза механизмов управления (правил принятия решения выделенным игроком — центром) для широкого класса ситуаций. Этим математическим аппаратом исследователи пользуются в ряде важных прикладных дисциплин — теории аукционов и теории контрактов, а также теории общественного выбора, политологии и др.

Базовый принцип открытого управления (принцип выявления, revelation principle), позволяю-

щий превратить синтез механизма из задачи на связанный экстремум в задачу оптимизации с ограничениями типа неравенств, был разработан в 1970-е гг. практически одновременно в СССР (В.Н. Бурков и его научная школа [18] и за рубежом (R. Myerson [12, 13] и др.). За последние годы в рамках этого научного направления был разработан комплекс оптимальных базовых механизмов управления для всех этапов управленческого цикла организации [10]. Каждый из этих механизмов оптимален — максимизирует выигрыш центра — и при этом обладает свойством неманипулируемости — невыгодности искажения сообщаемой информации для остальных игроков, агентов.

В практике организационного управления механизмы почти всегда применяются совместно, и важный открытый вопрос теории состоит в возможности сохранения положительных качеств механизмов при их комбинировании/«комплексировании». Поэтому актуальна задача выявления общих условий, которым должны удовлетворять механизмы управления, устойчивые к комплексированию. Удовлетворяющие этим условиям частные механизмы можно использовать как «кирпичики» при построении комплексного механизма управления, что позволит вывести теорию на качественно новый уровень и кардинально упростить применение ее результатов на практике.

¹ Работа выполнена при финансовой поддержке РНФ (грант № 16-19-10609).



С другой стороны, современный менеджмент предлагает для решения проблем организации набор «лучших практик» — хорошо показавших себя в отдельных случаях управленческих находок. Отбор «лучших практик», для конкретной организации, их адаптация и внедрение происходят неформально на основе интуиции и опыта руководителя и/или консультанта. Теоретические наработки, в которых применяются строгие формальные методы математической теории принятия решений, существуют лишь для решения отдельных базовых задач организационного управления, будь то распределение делимого или неделимого ресурса, выработка коллективного решения, снижение сроков реализации проекта или составления плана производств. Однако в практике организационного управления механизмы нужно применять совместно, и открытым вопросом теории остается возможность сохранения положительных качеств оптимальных механизмов управления при их комплексировании. Например, сохраняется ли оптимальность и неманипулируемость механизмов оптимального распределения ресурса и стимулирования при их последовательном применении в рамках одного бизнес-процесса? Общие условия на механизмы управления, устойчивые к комплексированию, в форме дополнительных ограничений на задачу синтеза базового механизма должны позволить использовать получаемые оптимальные базовые механизмы как «кирпичики» при построении комплексного механизма управления, избавляя разработчика от риска потери хороших свойств механизмов и от необходимости трудоемкого анализа комплексного механизма управления.

«Двойственной» к проблеме комплексирования является проблема декомпозиции механизмов управления — возможности представления некоторого механизма в виде совокупности нескольких взаимосвязанных более простых механизмов. Основаниями декомпозиции могут быть или иерархические уровни системы, или периоды времени и др. Вопросы о наследовании свойств в процессе декомпозиции возникают те же, что и в случае комплексирования. Далее основное внимание уделяется комплексированию.

Изложение материала настоящей работы, посвященной постановке задачи комплексирования механизмов управления организационно-техническими системами, имеет следующую структуру. Сначала приводится общее описание механизмов управления (§ 1) и задачи их комплексирования (§ 2). Затем проблема комплексирования анализируется с теоретико-игровой точки зрения (§ 3), после чего рассматриваются два типовых способа — параллельного и последовательного — комплексирования (§ 4 и 5 соответственно). Обсуждению проблем манипулируемости комплексных меха-

низмов посвящен § 6. В заключении представлены основные результаты и указаны перспективы дальнейших исследований.

1. ЭЛЕМЕНТАРНЫЕ И КОМПЛЕКСНЫЕ МЕХАНИЗМЫ УПРАВЛЕНИЯ

Общее определение *механизма* в Словаре русского языка С.И. Ожегова — «система, устройство, определяющее порядок какого-либо вида деятельности». Применительно к организационным системам *механизм* функционирования — это совокупность правил, законов и процедур, регламентирующих взаимодействие участников системы; *механизм управления* — совокупность процедур принятия управленческих решений [10].

Рассмотрим базовую модель организационной или организационно-технической системы (ОТС), состоящей из управляющего органа — центра (Ц) — и управляемого субъекта/объекта — агента (А) — рис. 1 [8, 10, 11, 14]. Имеются: на входе — управляемого субъекта (состояние управляемой системы). Обратная связь обеспечивает центр информацией о состоянии управляемой системы. В данном случае механизм управления совпадает с управлением (в общем случае механизм управления — отображение множества действий агента во множество действий центра — см. далее и работы [11, 15]).

Пусть состояние агента определяется его действием $y \in A$, принадлежащим множеству допустимых действий A. Допустим, управление $u \in U$, где U — множество допустимых управлений U. Пусть также задан κ ритерий эффективности функционирования системы $\Phi(u, y)$: $A \times U \to R^1$, который зависит от переменных, описывающих эту систему, т. е. от управления и от состояния системы.

Предположим, что известна реакция управляемого субъекта на то или иное управление. Простейший вид такой реакции — когда состояние субъекта является известной функцией от управления: y = G(u), где $G(\cdot)$ — модель активного управляемого субъекта (или объекта управления в случае,

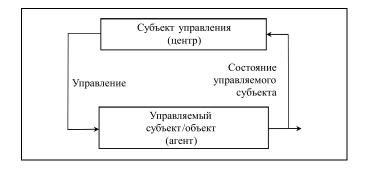


Рис. 1. Базовая модель организационной системы



когда объектом управления служит пассивная техническая система), которая описывает его реакцию на управляющее воздействие. Если подставить $G(\cdot)$ в критерий эффективности функционирования, то получим функционал $K(u) = \Phi(u, G(u))$, который будет зависеть только от управления. Этот функционал называется эффективностью управления.

Задача управления заключается в поиске оптимального управления, т. е. допустимого управления $u \in U$, обладающего максимальной эффективностью: $K(u) \to \max_{u \in U} M$

ность вычисляется как $K^* = \max_{u \in U} \Phi(u, G(u))$.

Отметим, что приведенная постановка простейшая — в общем случае механизм управления включает в себя этапы сообщения информации (здесь возникают задачи обеспечения неманипулируемости — выгодности для агентов сообщения достоверной информации), планирования (здесь возникают задачи обеспечения согласованности управлений — выгодности для агентов выполнения планов) и реализации [6, 8—10].

Элементарным (или базовым) механизмом назовем механизм, не включающий в себя другие механизмы (не допускающий дальнейшей декомпозиции в рамках принятой степени детализации). Комплексный механизм — механизм, включающий в себя один или несколько элементарных или других комплексных механизмов. Соответственно, комплексирование — процесс построения комплексного механизма, декомпозиция — выделение для некоторого комплексного механизма более простых его составляющих. В процессах комплексирования/декомпозиции одними из ключевых являются вопросы полноты набора элементарных механизмов — обеспечивает ли их совокупность возможность построения любого комплексного механизма из заданного множества, и их минимальности — каков минимальный полный набор элементарных механизмов.

Простейшая (базовая) модель ОТС включает в себя один управляемый субъект — агент — и один управляющий орган — центр, которые принимают решения однократно и в условиях полной информированности.

Перечислим расширения базовой модели [11].

- 1. Динамические *OTC* (в которых участники принимают решения многократно расширение по предмету управления «порядок функционирования»).
- 2. *Многоэлементные ОТС* (в которых имеется несколько агентов, принимающих решения одновременно и независимо, расширение по предмету управления «состав»).

- 3. *Многоуровневые ОТС* (имеющие трех- и более уровневую иерархическую структуру расширение по предмету управления «структура»).
- 4. *OTC с распределенным контролем* (в которых имеется несколько центров, управляющих одними и теми же агентами расширение по предмету управления «структура»).
- 5. *OTC с неопределенностью* (в которых участники не полностью информированы о существенных параметрах расширение по предмету управления «информированность»).
- 6. ОТС с ограничениями совместной деятельностии (в которых существуют глобальные ограничения на совместный выбор агентами своих действий расширение по предмету управления «множества допустимых стратегий»).
- 7. *OTC с сообщением информации* (в которых одно из действий агентов заключается в сообщении информации друг другу и/или центру расширение по предмету управления «множества допустимых стратегий»).

Таким образом, если в базовой модели используется некоторый элементарный механизм, то переход к тому или иному из первых четырех расширений базовой модели (каждое из которых может рассматриваться как совокупность взаимосвязанных базовых моделей) порождает проблему комплексирования элементарных механизмов — построения нового комплексного механизма, представляемого в виде совокупности взаимосвязанных элементарных механизмов.

Еще одним расширением модели представляется целесообразным считать рассмотрение комплексного механизма, декомпозируемого на набор базовых — ОТС с комплексным механизмом. Однако, в отличие от перечисленных выше расширений, данное расширение не может применяться к базовой модели ОТС, его применение целесообразно рассматривать с одним из первых четырех расширений базовой модели (каждое из которых может рассматриваться как совокупность взаимосвязанных базовых моделей. В смысле увеличения числа агентов и числа механизмов для систематизации результатов современной теории управления организационными системами целесообразно рассматривать классификацию, представленную в таблице.

В случае многоэлементных ОТС, в которых агенты принимают решения одновременно, имеет смысл условно говорить о параллельном комплексировании. В случае динамических ОТС, многоуровневых ОТС или ОТС с распределенным контролем — о последовательном комплексировании, где последовательность «соединения» элементарных механизмов определяется соответственно временем и порядком принятия решений участниками ОТС. Формальные модели параллельного



и последовательного комплексирований рассматриваются далее. В общем случае можно говорить о сетевом комплексировании, при котором каузальная последовательность «взаимодействия» механизмов описывается сетью, а такты времени соответствуют фронту в сети. Принимая во внимание постулат о том, что одновременное взаимодействие одного центра с одним агентом по нескольким механизмам невозможно, можно указать, что расширения 1.1, 2.2 и 3.2 допускают только последовательное комплексирование, в то время как исключительно параллельное комплексирование возможно в расширениях 2.1 и 3.1. При этом в расширениях 2.1., 2.3., 3.1 и 3.3 возможно сетевое комплексирование.

Кроме того, важным признаком классификации следует считать степень взаимосвязанности базовых механизмов, входящих в комплексный. По аналогии с классификацией отдельной базовой задачи управления — задачи стимулирования [9] введем понятие сильной и слабой связанности между базовыми механизмами управления и несвязанности.

Будем называть два механизма управления *сла- босвязанными*, если решение, принимаемое в рам-

ках одного механизма управления, влияет лишь на ограничения, накладываемые на действия агентов в другом механизме, и наоборот. Например, решение, принимаемое в рамках механизма распределения ресурсов [10], влияет лишь на то, какой размер вознаграждения доступен центру для вознаграждения отдельного агента при решении задачи стимулирования [10].

Будем называть два механизма *сильносвязанны- ми* в том случае, если решение, принимаемое в рамках одного из механизмов, влияет на параметры другого механизма. Например, решение, которое принимается в рамках механизма активной экспертизы [10] при оценивании важности отдельного проекта, влияет на вес данного проекта при распределении ресурсов между несколькими проектами на основе приоритетного механизма распределения ресурсов [10].

Соответственно, *несвязанными* будут считаться механизмы, принятие решений в рамках которых никак не взаимосвязанно.

В случае комплексных механизмов основная задача состоит в оценке их эффективности (в том числе в сравнении с эффективностями соответствующих элементарных механизмов) и выборе спо-

Расширения квалификации для комплексных механизмов

і асширения квалификации для комплексивы механизмов					
Состав ОТС	Комплекс механизмов				
Один центр, один агент	1.1. Динамические ОТС с одним агентом и комплексным механизмом. Центр с одним агентом реализует сложное взаимодействие, что можно рассматривать как динамическую ОТС, в которой на разных этапах центр и агент взаимодействуют в рамках разных механизмов управления				
Один центр, несколько агентов	2.1. <i>Многоэлементные ОТС с комплексным механизмом, декомпозируемым по агентам</i> . Центр взаимодействует в рамках каждого базового механизма или «подкомплекса» механизмов, из которых составлен комплекс, только с одним агентом (или непересекающимися подмножествами агентов), в ОТС образуется «цепочка» подсистем				
	2.2. <i>Многоэлементные ОТС с комплексным механизмом, декомпозируемым по механизмам</i> . Центр реализует одинаковое сложное взаимодействие с каждым из агентов. Сохраняется «классическая» веерная структура ОТС				
	2.3. <i>Многоэлементные ОТС с комплексным механизмом</i> . Сложная структура пересечений подмножеств агентов для каждого из базовых механизмов				
	3.1. <i>Многоуровневые ОТС с комплексным механизмом</i> . При рассмотрении базовых механизмов, входящих в комплексный, наблюдается «делегирование» полномочий: в рамках отдельных базовых механизмов отдельные агенты становятся центрами промежуточного уровня и им выделяются свои подчиненные из подмножества агентов. При этом на разных уровнях иерархии возможны разные ситуации из других разделов данной классификации				
Несколько центров, один агент	4.1. ОТС с распределенным контролем и с комплексным механизмом, декомпозируемым по центрам. В рамках каждого базового механизма с агентом взаимодействует один центр или непересекающееся подмножество центров				
	4.2. <i>OTC с распределенным контролем и с комплексным механизмом, декомпозируемым по механизмам</i> Все центры реализует одинаковое сложное взаимодействие с каждым из агентов. Сохраняется «классическая» веерная структура OC				
	4.3. <i>OTC с распределенным контролем и с комплексным механизмом.</i> Сложная структура пересечений подмножеств центров для каждого из базовых механизмов				



собов комплексирования, позволяющих синтезировать комплексный механизм максимальной эффективности. Также для комплексных механизмов требуют решения ряд общих проблем:

- допустимость (удовлетворение системе ограничений);
- непротиворечивость (достаточность информации, «согласованность» последовательных входов и выходов, ацикличность процедур принятия решений);
 - полнота и минимальность (см. выше);
- операциональность (возможность решения для комплексного механизма задачи синтеза, желательно аналитического);
- наследование (устойчивость) свойств механизмов (эффективность, неманипулируемость, согласованность и др.) по отношению к комплексированию/декомпозиции.

2. ОБЩЕЕ ОПИСАНИЕ ЗАДАЧИ КОМПЛЕКСИРОВАНИЯ МЕХАНИЗМОВ УПРАВЛЕНИЯ

Традиционно в теории управления социально-экономическими системами рассматривается система вложенных задач управления (решения более «частных» задач используются при решении более «общих»). На сегодня существуют два общих подхода к описанию модели ОТС, постановке и решению задач управления — «снизу вверх» и «сверху вниз» [11].

При первом подходе («снизу вверх») сначала решают частные задачи, а затем общие, используя полученные решения частных задач. Например, частной задачей может быть разработка системы мотивации. Если она решена для любого состава участников ОТС, то можно ставить задачу оптимизации состава — выбора такого состава, эффективность которого (при соответствующей оптимальной мотивации) максимальна. Достоинство такого подхода состоит в его конструктивности, недостаток — в высокой сложности, так как число вариантов решения задачи верхнего уровня может быть очень велико, а для каждого такого варианта необходимо решить соответствующий набор частных подзадач.

Бороться с этим недостатком можно с помощью второго подхода («сверху вниз»), в рамках которого сначала решают задачи верхнего уровня, а полученные решения используют в качестве ограничений для решения более частных задач. Однако общих рецептов в рамках второго подхода на сегодня не получено.

Приведем, следуя работе [16], постановку задачи синтеза оптимального комплекса механизмов управления ОТС. Пусть ОТС описывается набором (вектором) множества $L = \{1, 2, ..., l\}$ переменных:

$$y = (y_1, y_2, ..., y_l) \in A' = \prod_{i \in L} A_i$$
, где $y_i \in A_i$, $i \in L$, и существуют глобальные ограничения A^* на комбинации переменных: $y \in A' \cap A^*$.

Под механизмом $u(\cdot) \in \Xi$ будем понимать отображение множества $M_u \subseteq L$ значений управляемых переменных во множество $K_u \subseteq L$ значений управляющих переменных, т. е. $u: A_{M_u} \to A_{K_u}$, где $A_{M_u} = \prod_{i \in M_u} A_i, \ A_{K_u} = \prod_{i \in K_u} A_i$. Будем считать, что множество Ξ допустимых механизмов таково, что

множество Ξ допустимых механизмов таково, что для любого механизма $u(\cdot) \in \Xi$ выполнены глобальные ограничения, т. е.

$$\begin{split} \Xi &= \{u(\cdot) | \forall (y_{M_u}, \ y_{K_u}) : y_{K_u} = \\ &= u(y_{M_u}) \to (y_{M_u}, \ y_{K_u}) \in \operatorname{Proj}_{M_u \cup K_u}(A^*) \}, \end{split} \tag{1}$$

где
$$y_{M_u} = (y_j)_{j \in M_u}, \ y_{K_u} = (y_j)_{j \in K_u}, \ \operatorname{Proj}_{M_u \cup K_u}(A^*)$$
 — проекция множества A^* на множество $M_u \cup K_u$.

Введем $\Sigma \subseteq \Xi$ — подмножество множества допустимых механизмов, $\Sigma \in 2^\Xi$ — множеству всех подмножеств множества Ξ . Обозначим Q_Σ — множество всевозможных последовательностей элементов множества, q_Σ — произвольный элемент множества Q_Σ .

Множество Σ механизмов назовем *непротиворечивым*, если

$$\exists q_{\Sigma} \in Q_{\Sigma}: \exists (u, ..., v) \in q_{\Sigma}: M_{u} \cap K_{v} \neq \emptyset. \quad (2)$$

Свойство непротиворечивости означает, что для данного набора механизмов не существует их последовательности, для которой нашлась бы переменная, которая была бы одновременно управляемой для первого механизма в этой последовательности и управляющей — для последнего. Отметим, что в случае последовательного (например, повторяющегося во времени) применения комплекса механизмов свойство (2) может нарушаться.

Непротиворечивость множества механизмов порождает в ОТС иерархию: множество параметров ОТС может быть упорядочено — на нижнем уровне находятся параметры из множества $L_1 = L \setminus \bigcup_{u \in \Sigma} K_u$, на следующем уровне — параметры, которые являются управляющими по отношению к параметрам нижнего уровня, но управляемыми для параметров, находящихся на более высоких уровнях иерархии, и т. д. (см. также теоретико-игровые модели синтеза иерархических структур в работе [15]).

Поставим в соответствие i-му параметру ОТС активного агента, обладающего целевой функцией f_i : $A \to \Re^1$, $i \in L_1$.



Предположим, что при заданном комплексе механизмов Σ агенты из множества L_1 будут стремиться выбирать равновесные по Нэшу стратегии (полностью по аналогии рассматриваются случаи использования других концепций решения игры агентов — гарантирующих стратегий, равновесий Байеса — Нэша и т. п.). Обозначим соответствующее множество равновесий Нэша как

$$\begin{split} E^N(\Sigma) &= \{ y_{L1} \in A_{L1} \mid \forall \ i \in L_1 \ \forall \ y_i \in A_i f_i(y_{L1}, \ u(y_{L1})) \geq \\ &\geq f_i(y_{L1} \ | y_i, \ u(y_{L1} | y_i)) \}, \end{split}$$

где $u(y_{L1})$ — действия, выбираемые агентами из множества $\bigcup_{u \in \Sigma} K_u$ (эти действия при заданном комплексе механизмов определяются действиями, выбираемыми агентами из множества L_1).

Пусть на множестве A' состояний системы задан функционал $\Phi(\cdot)$: $A' \to \Re^1$, характеризующий эффективность ее функционирования. Задача *синтеза оптимального комплекса механизмов* может формулироваться таким образом:

$$\min_{y_{L1} \in E^{N}(\Sigma)} \Phi(y_{L1}, \ u(y_{L1})) \to \max_{\Sigma \in 2^{\Xi}, (1), (2)}, \tag{3}$$

т. е. требуется найти непротиворечивый и удовлетворяющий глобальным ограничениям (условия (2) и (1) соответственно) комплекс механизмов, обладающий максимальной гарантированной эффективностью.

Отметим, что при формулировке задачи (3) мы не учитывали явным образом интересы агентов из множества $\bigcup_{u \in \Sigma} K_u$. Если предположить, что каждый из них может самостоятельно выбирать определенные механизмы управления, то получим задачу, аналогичную задаче структурного синтеза, описанной в работе [15].

Ситуация также осложняется тем фактом, что решение задачи синтеза отдельного механизма управления (и эффективность этого решения!), как правило, не устойчиво по параметрам модели [17, 18], что накладывает существенные ограничения на возможности применения методов первого и более всяких порядков при оптимизации «вложенных» механизмов.

На сегодня общие методы решения задачи (3) или задачи структурного синтеза неизвестны. Поэтому на практике при синтезе комплекса механизмов либо решают задачу последовательного синтеза, либо согласовывают в рамках той или иной метамодели отдельные оптимальные механизмы управления.

Применение подходов управления проектами и календарно-сетевого планирования, в частности, методов сетевого программирования [19], может оказаться продуктивным при исследовании комп-

лексных механизмов. Действительно, функционирование системы при заданном комплексном механизме можно представлять в виде сети без контуров, вершины которой соответствуют центрам преобразования информации или центрам производства, а дуги выражают информационные или технологические связи. Такое представление близко к описанию функционирования как бизнес-процесса с тем отличием, что в ряде вершин (активных вершин) решения принимает агент, имеющий свои интересы. Учитывая, что функционирование системы, как правило, разбивается на периоды, каждый период можно рассматривать как проект. Для оценки эффективности комплексного механизма необходимо получить оценку потенциала системы (максимальное значение целевой функции при полной информированности и выполнении всех планов). Сравнение оценки при заданной системе механизмов и принятых моделях поведения агентов с оценкой потенциала позволяет оценить эффективность комплексного механизма. Меняя отдельные блоки комплексного механизма, можно решать задачу оптимизации.

3. ИГРЫ И СТРУКТУРЫ

Как отмечалось в § 2 и в работах [11, 15], последовательность выбора действий участниками ОТС (порядок функционирования) тесно связана со структурой системы. Опишем кратко, следуя работе [11], эту взаимосвязь, переходя индуктивно от простых к более сложным моделям.

Базовой моделью принятия решений можно считать ситуацию принятия решений одним агентом в условиях полной информированности — см рис. 2, а. При этом считается выполненной так называемая гипотеза рационального поведения (ГРП) [8, 11], в соответствии с которой агент выбирает допустимое действие, максимизирующее его целевую функцию. Следующий этап «усложнения»

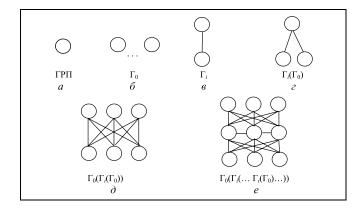


Рис. 2. Игры и структуры (i = 1, 2, 3) [20]: $a, \delta, s, \varepsilon, \partial$ и e - см. в тексте



состоит во взаимодействии нескольких агентов, действующих в соответствии с ГРП — uгpa, в которой агенты выбирают свои действия однократно, одновременно и независимо, т. е. игра Γ_0 в нормальной форме — см. рис. 2, δ . На следующем этапе осуществляется переход к иерархическим играм Γ_i , где i=1, 2 или 3 [11], в которых последовательность ходов фиксирована — первым делает ход центр, а затем — агент. Можно усложнять модель и дальше, переходя к все более сложным играм.

Например, если имеется структура «один центр — несколько агентов» (рис. 2, $\it e$), то взаимодействие агентов, находящихся на одном уровне, можно описывать игрой $\Gamma_{\it o}$; взаимодействие «центр — агент» описывается игрой $\Gamma_{\it i}$. Тогда такую структуру можно представить игрой $\Gamma_{\it i}$, определенной на игре $\Gamma_{\it o}$, условно обозначив ее $\Gamma_{\it i}(\Gamma_{\it o})$.

Далее, пусть имеется несколько (центров) и несколько агентов (рис. 2, ∂). На нижнем уровне агенты разыгрывают игру Γ_0 . «Над ними» центры разыгрывают иерархическую игру Γ_i , но центры, в свою очередь, разыгрывают на своем уровне игру Γ_0 . Итак, получили игру $\Gamma_0(\Gamma_i(\Gamma_0))$.

Если взять более сложную структуру (например, рис. 2, e), то ей будет соответствовать иерархическая игра между уровнями, и «обычная» игра в нормальной форме на каждом из уровней: $\Gamma_0(\Gamma_i(...\Gamma_i(\Gamma_0)...))$.

Основная идея «комплексирования игр» заключается в том, чтобы декомпозировать сложную организационную структуру (сложную игру) на набор более простых и воспользоваться результатами исследования последних. Это возможно, так как между играми и структурами существует глубокая связь — момент принятия субъектом решений определяет его «место» в организационной иерархии (см. подробности в работе [15]).

4. ПАРАЛЛЕЛЬНОЕ КОМПЛЕКСИРОВАНИЕ

Рассмотрим проблему параллельного комплексирования на примере построения комплексного механизма из двух элементарных механизмов при переходе от двух структур (см. рис. 2, θ) с одним и тем же центром к структуре, представленной на рис. 2, e.

Эффективности первого и второго элементарных механизмов обозначим соответственно

$$K_1^* = \max_{u_1 \in U_1} \Phi_1(u_1, G_1(u_1))$$

и
$$K_2^* = \max_{u_2 \in U_2} \Phi_2(u_2, G_2(u_2)).$$
 (4)

Обозначим $(y_1, y_2) = \hat{G}(u_1, u_2)$, тогда эффективность комплексного механизма определяется аналогично:

$$K^* = \max_{(u_1, u_2) \in U} \Phi(u_1, u_2, \hat{G}(u_1, u_2)).$$
 (5)

Ключевым при комплексировании является вопрос о том, как соотносятся эффективности K^* и (K_1^* , K_2^*)? Ответить на него невозможно, не получив ответы на следующие вопросы — как соотносятся:

- критерии эффективности $\Phi(\cdot)$ и $(\Phi_1(\cdot), \Phi_2(\cdot))$; возможные варианты критерий $\Phi(\cdot)$ монотонен по критериям $\Phi_1(\cdot)$ и $\Phi_2(\cdot)$ или совпадает с последним в цепочке критерием;
- множества допустимых управлений U и $(U_1,\ U_2)$; возможные варианты управления независимы $(U=U_1\times U_2)$, или существуют ресурсные ограничения $(0\le u_1\le C_1,\ 0\le u_2\le C_2\to 0\le u\le C_1+C_2)$, или управление в комплексном механизме должно быть унифицированным, т. е. одинаковым для всех агентов $(u=(v,\ v))$, и т. д.;
- модели поведения агентов $\hat{G}(\cdot)$ и $(G_1(\cdot))$ и $G_2(\cdot)$); возможные варианты агенты независимы $(\hat{G}(u_1, u_2) = (G_1(u_1), G_2(u_2))$ или зависимы.

Если критерий эффективности аддитивен, существуют ресурсные ограничения на управление, а агенты независимы, то, очевидно,

$$K^* \ge K_1^* + K_2^*.$$
(6)

Величина $\Delta = K^* - (K_1^* + K_2^*)$ может быть условно названа *ценой комплексирования*.

Пример 1. Рассмотрим приведенную в работе [11] модель стимулирования в ОТС со слабо связанными агентами. Пусть $N = \{1, 2, ..., n\}$ — множество агентов, $y_i \ge 0$ — действие i-го агента, $c_i(y_i)$ — затраты i-го агента (монотонно возрастающая выпуклая функция), $i \in N$. Если $u_i(y_i)$ — система стимулирования, используемая центром, то целевая функция агента $f_i(u_i, y_i) = u_i(y_i) - c_i(y_i)$, а модель поведения агента имеет вид:

$$G_i(u_i, y_i) = \arg\max_{y_i \ge 0} f_i(u_i, y_i).$$

Пусть размеры индивидуальных вознаграждений агентов ограничены заданными значениями $\{C_i\}_{i\in N}$, т. е. $\forall y_i \geq 0 \ u_i(y_i) \leq C_i, \ i \in N$. Оптимальной является система стимулирования i-го агента [11]:

$$u_i(x_i, y_i) = \begin{cases} c_i(x_i), & y_i \ge x_i, \\ 0, & y_i < x_i, \end{cases}$$

где $x_i \in [0; c_i^{-1}(C_i)]$ — план *i*-го агента, при этом $G_i(u_i(x_i, y_i)) = x_i, i \in N$.



Предположим, что центр получает монотонно возрастающий линейный или вогнутый доход $H_i(y_i)$ от деятельности i-го агента, т. е. целевая функция центра $\Phi_i(u_i(x_i,\ y_i),\ y_i)=H_i(y_i)-u_i(x_i,\ y_i)$. Подставляя $G_i(u_i(x_i,\ y_i))=x_i$, получим $\Phi_i(u_i(x_i,\ y_i),\ y_i)=H_i(x_i)-c_i(x_i)$. Следовательно.

$$K_i^*(C_i) = \max_{x_i \in [0; c_i^{-1}(C_i)]} [H_i(x_i) - c_i(x_i)], \quad i \in \mathbb{N}.$$
 (7)

Рассмотрим теперь комплексный механизм, в котором центр управляет n агентами, используя для каждого из них систему стимулирования (4). Предположим, что фонд заработной платы (ФЗП) ограничен значением R, т. е.

$$\sum_{i \in N} C_i \leq R,$$

и центр имеет целевую функцию, аддитивную по критериям одноэлементных задач: $\Phi(u, y) = \sum_{i \in N} [H_i(y_i) - u_i(x_i, y_i)]$, где $y = (y_1, y_2, ..., y_n)$ — вектор действий агентов, $y \in A' = \prod_{i \in N} A_i$, $u = (u_1, u_2, ..., u_n)$ — вектор управлений. Тогда

$$K^{*}(R) = \max_{\{x_{i} \in [0; c_{i}^{-1}(C_{i})]\}, \{C_{i} \ge 0 | \sum_{i \ge N} C_{i} \le R\}} \sum_{i \in N} H_{i}(x_{i}) - R. \quad (8)$$

Легко видеть, что эффективности (7) и (8) удовлетворяют условию (6), т. е. $\forall \{\{C_i \geq 0\}_{i \in N} | \sum_{i \in N} C_i \leq R\}$ выполнено

 $K^* \ge \sum_{i \in N} K_i^*(C_i)$, следовательно, цена комплексирования неотрицательна.

Если ФЗП является переменной величиной, то его оптимальное значение R^* может быть найдено как решение следующей задачи: $R^* = \underset{R \geq 0}{\operatorname{arg\,max}} K^*(R)$. Отметим, что поиск оптимального размера ФЗП может рассматриваться уже как задача последовательного комплексирования. \blacklozenge

Завершив рассмотрение примера, отметим, что при декомпозиции механизмов сформулированные выше вопросы звучат и решаются обычно несколько проще. Например, пусть имеется комплексный механизм с эффективностью (5). Тогда при параллельной декомпозиции можно определять эффективности первого и второго частных механизмов через критерий комплексного механизма:

$$K_1^*(u_2) = \max_{u_1 \in U_1} \Phi(u_1, u_2, \text{Proj}_1(\hat{G}(u_1, u_2)))$$

и
$$K_2^*(u_1) = \max_{u_2 \in U_2} \Phi(u_1, u_2, \operatorname{Proj}_2(\hat{G}(u_1, u_2))),$$

считая, например, что $U_1 = \operatorname{Proj}_1(U)$, $U_2 = \operatorname{Proj}_2(U)$, но помня, что покоординатная оптимизация может привести к недопустимому решению как по состоянию агентов, так и по управлению (так как может быть $\hat{G}(u_1, u_2) \neq (G_1(u_1), G_2(u_2))$ и/или $U \neq \operatorname{Proj}_1(U) \times \operatorname{Proj}_2(U)$).

5. ПОСЛЕДОВАТЕЛЬНОЕ КОМПЛЕКСИРОВАНИЕ

Рассмотрим проблему последовательного комплексирования на примере построения комплексного механизма из двух элементарных механизмов. При последовательном комплексировании эффективность первого механизма определяется, как и в выражении (4), т. е. $K_1^* = \max_{u_1 \in U_1} \Phi_1(u_1, G_1(u_1))$, а

эффективность второго механизма, а также множество соответствующих допустимых управлений и модель поведения второго агента, в общем случае зависит от управления, использованного в первом механизме

$$K_2^*(u_1) = \max_{u_2 \in U_2(u_1)} \Phi_2(u_2, G_2(u_1, u_2)).$$
 (9)

Эффективность комплексного механизма определим как

$$K^* = \max_{(u_1, u_2) \in U} \Phi(u_1, u_2, \hat{G}_1(u_1), \hat{G}_2(u_1, u_2)). \quad (10)$$

Соотношение эффективностей K^* и ($K_1^*, K_2^*(u_1)$), как и при параллельном комплексировании, в рассматриваемом случае зависит от того, как соотносятся соответствующие критерии эффективности, множества допустимых управлений и модели поведения агентов (см. выше). Что касается эффективности как при последовательной декомпозиции, так и при последовательном комплексировании, естественным представляется считать критерием терминальное значение эффективности, т. е. соответствующее последнему шагу («моменту времени») — см. выражения (9) и (10). В прикладных «экономических» задач управления критерием эффективности элементарного механизма управления, соответствующего некоторому этапу последовательности, может быть та «добавленная стоимость» (в терминах терминального критерия), которая создается на данном этапе.

В случае аддитивных критериев эффективности (вариант, соответствующий переходу от статичес-



кой к динамической модели ОТС) получаем, что выражение (10) принимает вид:

$$\begin{split} K^* &= \max_{u_1 \in U_1} \max_{u_2 \in U_2(u_1)} \{ \Phi_1(u_1, \ G_1(u_1)) \ + \\ &+ \Phi_2(u_2, \ G_2(u_1, \ u_2))]. \end{split}$$

Получаем аналог принципа оптимальности Беллмана:

$$K^* = K_1^* + \max_{u_1 \in U_1} K_2^*(u_1).$$

Очевидно, что $\forall u_1 \in U_1 K^* \geq K_1^* + K_2^*(u_1)$.

6. ПРОБЛЕМА НЕМАНИПУЛИРУЕМОСТИ КОМПЛЕКСНЫХ МЕХАНИЗМОВ

В § 4 и 5 мы рассматривали соотношение эффективностей элементарных и соответствующих комплексных механизмов. Аналогичные вопросы можно задать и касательно свойства неманипулируемости:

- 1) будет ли неманипулируемым комплексный механизм, построенный из совокупности неманипулируемых механизмов?
- 2) для отдельных задач управления показано, что их эффективное решение достаточно искать в классе неманипулируемых механизмов, если комплексная задача управления состоит из подобных «элементарных» задач, то достаточно ли ее решение искать в классе механизмов, «собранных» из неманипулируемых элементарных механизмов?

Ключ к ответу на вопросы, связанные с неманипулируемостью, кроется в том, насколько сильно может быть декомпозирована комплексная задача управления в следующих аспектах. Прежде всего, насколько сильна связь между базовыми механизмами, входящими в комплексный. Далее, насколько взаимосвязаны интересы агентов в рамках отдельных механизмов, как и как это формализуется в форме целевых функций агентов. Суть в том, что проблему манипулируемости для комплексного механизма можно рассматривать как проблему манипулируемости для задачи планирования в многомерном пространстве — для каждого из механизмов планирование осуществляется в своем подпространстве параметров. А для данного класса задач показано (см., например, [21, 22]), что неманипулируемые механизмы существуют, если интересы агентов формализуются многомерно-однопиковыми предпочтениями, в которых наилучшее значение по любому из планируемых параметров [23] не зависит от значений других планируемых параметров. При этом неманипулируемые механизмы не должны быть сильносвязанными. Ответ на то, достаточно ли искать эффективное решение в классе неманипулируемых механизмов положителен лишь для ситуаций, когда каждый из агентов заинтересован не более, чем в одном параметре планирования [22, 24].

Поэтому положительный ответ на первый вопрос, по всей видимости, может быть получен в случае слабой связанности элементарных механизмов, входящих в комплексный, или когда комплексный механизм декомпозируем по агентам (расширение 2.1), на второй вопрос — когда комплексный механизм декомпозируем по агентам.

Наконец, комплементарным вопросу неманипулируемости является вопрос эквивалентности механизмов [9], который для комплексных механизмов может быть сформулирован следующим образом. Пусть имеется двухуровневая ОТС с некоторым механизмом планирования. Вопрос заключается в том, существуют ли трехуровневая ОТС (с тем же составом агентов) и механизм планирования (комплексный) в ней такие, чтобы равновесные сообщения и назначаемые планы в этих ОТС были одинаковы. Эту задачу называют задачей идеального агрегирования в механизмах планирования — см. [25].

ЗАКЛЮЧЕНИЕ

Отсутствие общих методов аналитического решения задачи синтеза оптимального комплекса механизмов управления организационными системами приводит к необходимости систематизации и выделения соответствующих частных случаев. Возможные основания для этого приведены в настоящей работе.

Глобальная задача заключается в выявлении максимально общих условий, которым должны удовлетворять механизмы управления, устойчивые к комплексированию. Такие условия можно сравнить с условиями декомпозиции задач оптимизации, позволяющими свести решение одной сложной задачи к решению набора более простых (подобные подходы развиваются достаточно давно в рамках теории оптимизации [26—30], теории согласования иерархических решений [6, 14, 31, 32], иерархическом моделировании [20, 32] и др., а в настоящее время особенно интенсивно в рамках распределенной оптимизации [34, 35] и др. — см. обзоры в работах [20, 36].

Действительно, проблема выявления ограничений, порождаемых комплексированием, сродни декомпозиции решения оптимизационной задачи (в целях распараллеливания), а также задаче об идеальном агрегировании, однако здесь объединяемые «куски» (механизмы) существенно разнородны, а связи между отдельными механизмами включают не только общий критерий оптимальности, но и «технологические» связи через общие переменные, фиксируемые одним механизмом, а



используемые другим, ограничения на общие ресурсы и др.

Перспективным представляется создание теории для анализа и синтеза комплексных механизмов организационного управления на базе разрабатываемых теоретико-игровых и оптимизационных моделей и методов. Наличие библиотеки базовых механизмов управления и правил их эффективного комплексирования позволит кардинально упростить построение комплексных механизмов организационного управления и существенно приблизить теорию механизмов управления к нуждам современного менеджмента. С практической точки зрения, соответствующие теоретические результаты могут стать основой прикладной методики комплексного реинжиниринга механизмов управления производственных и коммерческих компаний, органов местной и государственной власти и т. д.

ЛИТЕРАТУРА

- Fudenberg D., Tirole J. Game Theory. Cambridge: MIT Press, 1995. — 579 c.
- Myerson R. Game Theory: Analysis of Conflict. London: Harvard Univ. Press, 1991. — 568 p.
- Algorithmic Game Theory / Ed. N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani). — N.-Y.: Cambridge University Press, 2009. — 776 p.
- 4. *Mansour Y.* Computational Game Theory. Tel Aviv: Tel Aviv University, 2003. 150 p.
- Germeier Yu. Non-antagonistic Games. Dordrecht, Boston:
 D. Reidel Pub. Co., 1986. 331 p.
- 6. *Бурков В.Н., Кондратьев В.В.* Механизмы функционирования организационных систем. М.: Наука, 1981. 384 с.
- 7. *Bolton P., Dewatripont M.* Contract Theory. Cambridge: MIT Press, 2005. 740 p.
- 8. *Бурков В.Н.* Основы математической теории активных систем. М.: Наука, 1977. 255 с.
- Burkov V., Goubko M., Korgin N., Novikov D. Introduction to Theory of Control in Organizations. — N.-Y.: CRC Press, 2015. — 352 p.
- Mechanism Design and Management: Mathematical Methods for Smart Organizations / Ed. prof. D. Novikov. — N.-Y.: Nova Science Publishers, 2013. — 163 p.
- Novikov D. Theory of Control in Organizations. N.-Y.: Nova Science Publishers, 2013. — 341 p.
- Myerson R. Incentive Compatibility and the Bargaining Problem // Econometrica. 1979. — Vol. 47, N 1. — P. 61—74.
- Myerson R. Optimal Coordination Mechanisms in Generalized Principal-Agent Problems // Journal of Mathematical Economy. — 1982. — Vol. 10, N 1. — P. 67—81.
- 14. *Ашимов А.А., Бурков В.Н., Джапаров Б.А., Кондратьев В.В.* Согласованное управление активными производственными системами. М.: Наука, 1986. 248 с.
- Новиков Д.А. Сетевые структуры и организационные системы. М.: ИПУ РАН, 2003. 102 с.
- Новиков Д.А., Иващенко А.А. Модели и методы организационного управления инновационным развитием фирмы. — М.: КомКнига, 2006. — 332 с.
- Novikov D.A. Management of Active Systems: Stability or Efficiency // Systems Science. 2001. Vol. 26, N 2. P. 85—93.

- Novikov D.A. Stability of Solutions and Adequacy of the Determinate Incentive Models in Active Systems // Automation and Remote Control. 1999. Vol. 60, N 7. P. 999—1004.
- 19. *Burkov V.N., Burkova I.V.* Network programming technique in project management problems // Automation and Remote Control. 2012. Vol. 73, N 7. P. 1242—1255.
- Novikov D. Hierarchical Models in Modern Control Theory / Challenges in Automation, Robotics and Measurement Techniques. — Heidelberg: Springer, 2016. — P. 3—12.
- 21. *Bondarik V.N., Korgin N.A.* Resource allocation mechanisms based on strategy-proof symmetrical anonymous voting procedures with delegation // Automation and Remote Control. 2013. Vol. 74, N 9. P. 1557—1566.
- 22. Korgin N.A. Representing a sequential allotment rule in the form of a strategy-proof mechanism of multicriteria active expertise // Automation and Remote Control. 2014. Vol. 75, N 5. P 983—995
- 23. *Barberá S.* Strategyproof Social Choice // Handbook of social choice and welfare. 2011. Vol. 2. P. 731—831.
- Moulin H. One dimensional mechanism design // Web and Internet Economics: 11th International Conference, WINE 2015, Amsterdam, The Netherlands, December 9–12, 2015, Proceedings. Springer, 2015. Vol. 9470. P. 436.
- 25. *Новиков Д.А.* Механизмы функционирования многоуровневых организационных систем. М.: Фонд «Проблемы управления», 1999. 50 с.
- 26. *Павловский Ю.Н., Смирнова Т.Г.* Проблема декомпозиции в математическом моделировании. М.: ФАЗИС, 1998. 266 с.
- 27. *Первозванский А.А., Гайцгори В.Г.* Декомпозиция, агрегирование и приближенная оптимизация. М.: Физматлит, 1979. 344 с.
- 28. *Цурков В.И.* Декомпозиция в задачах большой размерности. М.: Физматлит, 1981. 352 с.
- 29. *Lasdon L.* Optimization Theory for Large Systems. N.-Y.: MacMillan, 1970. 523 p.
- 30. *Nesterov Yu.* Introductory Lectures on Convex Optimization: a Basic Course. Heidelberg: Springer, 2014. 236 p.
- Ватель И.А., Ерешко Ф.И. Математика конфликта и сотрудничества. М.: Знание, 1973. 64 с.
- 32. Горелик В.А., Горелов М.А., Кононенко А.Ф. Анализ конфликтных ситуаций в системах управления. М.: Радио и связь, 1991.-288 с.
- 33. *Месарович М., Мако Д., Такахара И.* Теория иерархических многоуровневых систем. М.: Мир, 1973. 344 с.
- 34. Boyd S., Parikh N., Chu E., et al. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers // Foundations and Trends in Machine Learning. — 2011. — N 3 (1). — P. 1—122.
- 35. *Ren W., Yongcan C.* Distributed Coordination of Multi-agent Networks. London: Springer, 2011. 307 p.
- 36. *Novikov D.* Cybernetics: from Past to Future. Heidelberg: Springer, 2016. 107 p.

Статья представлена к публикации руководителем *PPC A.A.* Ворониным.

Бурков Владимир Николаевич — д-р техн. наук, зав. лабораторией, Институт проблем управления РАН им. В.А. Трапезникова, г. Москва, ⊠ vlab17@bk.ru,

Коргин Николай Андреевич — д-р техн. наук, вед. науч. сотрудник, Институт проблем управления РАН им. В.А. Трапезникова, г. Москва; профессор, Московский физико-технический институт (государственный университет), ⋈ nkorgin@ipu.ru,

Новиков Дмитрий Александрович — чл.-корр. РАН, зам. директора, Институт проблем управления РАН им. В.А. Трапезникова, г. Москва; зав. кафедрой, Московский физико-технический институт (государственный университет), ⊠ novikov@ipu.ru.



УДК 519.21:681.518

МОДЕЛЬ УПРАВЛЕНИЯ ЧЕЛОВЕЧЕСКИМ КАПИТАЛОМ ФИРМЫ

М.В. Белов

Предложена модель экономического управления человеческим капиталом предприятия, интенсивно использующего знания. Отмечено, что использование в модели базовых соотношений управленческого учета предприятия существенно повышает ее практическую применимость. На примере конкретной фирмы продемонстрировано решение задачи оптимизации затрат на человеческий капитал — управление численностью и развитием сотрудников в условиях неравномерной и неопределенной потребности в человеческих ресурсах со стороны бизнеса.

Ключевые слова: управление, человеческий капитал, знания, предприятие, жизненный цикл сотрудника, стохастическое моделирование.

ВВЕДЕНИЕ

Сотрудники всегда составляли и будут составлять основной актив практически любого бизнеса. На современном этапе развития общества, который многие исследователи и практики называют «экономикой знаний», ключевое значение приобретают вопросы развития сотрудников, управления знаниями и информацией, организационными структурами, объединяемые понятием «управление человеческим капиталом» [1—5].

Как объект исследований обосабливаются фирмы, интенсивно использующие знания (knowledge intensive firms), в которых основная доля добавленной стоимости формируется в виде знаний и благодаря использованию знаний сотрудников [6]. В настоящей работе будем рассматривать предприятие , интенсивно использующее знания и человеческий капитал (ПЗЧК). К ПЗЧК будем относить предприятия, формирование добавленной стоимости которых основано на создании и преобразовании информации, знаний. Это, например, фирмы, оказывающие профессиональные (финансовые, юридические, инженерные и другие) услу-

ги, банки, интернет-компании и другие экономические агенты, отличительные особенности бизнеса которых заключаются в ключевой роли знаний, навыков, эффективности сотрудников, высокой значимости технологических знаний и низкой капиталоемкости.

Тема управления человеческим капиталом не нова, к этой категории обращались экономисты еще в XIX—XX вв., но первым употреблением термина в современной экономической литературе считается работа [7].

К настоящему моменту практически во всей сфере управления человеческими ресурсами, кадрами, талантами, человеческим капиталом применяются качественные подходы, характерные для социологии, социальной психологии и смежных дисциплин. Рассматриваются и развиваются различные концепции удовлетворенности, лояльности, преданности, приверженности сотрудников, например, в работе [10]. Понятие «вовлеченности сотрудников» впервые введено в работе [11] и широко применяется сейчас, разработано множество методик оценки вовлеченности, в частности, [12]. Теория вовлеченности рассматривает вопрос о том, как компания может достичь своих стратегических целей, создавая такие условия для развития персонала, где каждый сотрудник, менеджер и руководитель делают все возможное для блага компании [10].

Достаточно полный обзор современных подходов и методов в области управления знаниями, талантами, человеческим капиталом выполнен в работе [13].

¹ Под предприятием понимаем [8, 9] комплексную социально-техническую систему, объединяющую взаимозависимые ресурсы (персонал, финансы, материалы, оборудование, здания и сооружения и др.), процессы, данные/знания и технологии, которые взаимодействуют друг с другом и с окружающей средой для достижения общих целей. Предприятиями являются фирмы или их объединения, подразделения фирм, рабочие группы, проектные команды.



В работе [14], посвященной управлению персоналом в вооруженных силах США, вводится формальная классификация систем развития персонала в организациях на основании двух основных признаков.

- Все организации относятся к «закрытому» или «открытому» типам. В первом случае вхождение нового сотрудника в организацию возможно только на нижние позиции, так как требуемая квалификация может быть приобретена только в данной организации, во втором случае на любые уровни.
- Наличие в организации формальной процедуры увольнения сотрудников в случаях, кроме невыполнения обязанностей, стимулирующей определенную сменяемость персонала соответственно общей стратегии организации.

На основании этих признаков выделяются четыре типа организаций [14]:

- «up-or-out»: вхождение в организацию только на начальные позиции и принудительное увольнение, следуя определенным процедурам и критериям отбора, отвечающим стратегическим целям организации;
- «up-and-stay»: вхождение в организацию только на начальные позиции и работа в течение всей карьеры, даже в отсутствии роста;
- «in-and-out»: вхождение в организацию на любые позиции в сочетании с принудительным или естественным увольнением;
 - «mixed»: сочетание предыдущих трех типов.

Разрабатываемая тема находится на пересечении как качественных, перечисленных выше, так и количественных подходов и методов — управленческого учета, исследования операций, стохастического моделирования, экономической теории фирмы и других смежных областей.

Концепции и практика управленческого учета применительно к человеческим ресурсам фирмы хорошо разработаны и изложены в большом числе работ, начиная с базовой монографии Е. Flamholtz [15] и других его работ [16], где излагаются подробные модели учета фактически понесенных затрат на персонал, включая не только заработную плату, но также затраты на найм, адаптацию, обучение сотрудников и др.

Работы в области организационной экономики анализируют численность и организационную структуру фирм, связи стратегии с организационными структурами и человеческим капиталом [17].

Данные модели хорошо описывают фактически произошедшие бизнес-явления, однако мало помогают при стратегическом планировании и оперативном управлении.

Нельзя не упомянуть признанные фундаментальными работы [18, 19] Г. Беккера, получившего

Нобелевскую премию за развитие теории человеческого капитала. Человеческий капитал определяется как имеющийся у каждого запас знаний, навыков, мотиваций. Инвестициями в него могут быть образование, накопление производственного опыта, охрана здоровья, географическая мобильность, поиск информации. При вкладывании средств в профессиональную подготовку и образование обучаемые и их родители ведут себя рационально, взвешивая соответствующие выгоды и издержки, они сопоставляют ожидаемую предельную норму отдачи от таких вложений с доходностью альтернативных инвестиций (процентами по банковским вкладам, дивидендами по ценным бумагам и др). В зависимости от того, что экономически целесообразнее, принимается решение либо о продолжении учебы, либо о ее прекращении. Нормы отдачи выступают, следовательно, как регулятор распределения инвестиций между различными типами и уровнями образования, а также между системой просвещения в целом и остальной экономикой. Высокие нормы отдачи свидетельствуют о недоинвестировании, низкие — о переинвестировании. Помимо теоретического обоснования Беккер первым осуществил и практический, статистически корректный подсчет экономической эффективности образования.

Несмотря на большое число фундаментальных и прикладных работ в рассматриваемой области, имеет место дефицит методов и подходов, обеспечивающих экономически, количественно, обоснованное принятие решений в области управления человеческим капиталом фирмы в практической деятельности, и в настоящей работе делается попытка восполнить этот пробел.

1. ПОСТАНОВКА ЗАДАЧИ

Предметом исследования сделаем экономические и количественные факторы управления человеческим капиталом ПЗЧК, используя известные результаты работ в смежных областях. Целью работы определим разработку модели управления человеческим капиталом (ЧК), предназначенную для формирования менеджментом компании обоснованных оптимальных решений.

Проанализируем основные требования к разрабатываемой модели.

Прежде всего, модель должна учитывать не только экономические аспекты жизненного цикла сотрудника в ходе работы на предприятии, но также рассматривать ЧК как составную часть социально-технической системы предприятия, т. е. отражать вторичность человеческого капитала как ресурса по отношению к потребности в нем со сто-



роны бизнеса, а следовательно, учитывать неопределенность бизнеса, выражающуюся в неравномерности потребности в ресурсах (ЧК) во времени. Общепринятым подходом к представлению неопределенных и неравномерных процессов служит стохастическое моделирование, которым воспользуемся в работе.

Человеческий капитал будем рассматривать как организованную в рамках ПЗЧК совокупность специалистов, обладающих определенными знаниями, навыками, квалификацией и мотивированных на решение ставящихся перед ними задач. Для формализации ЧК употребим понятие «функциональных домов» (ФД), введенное фирмой «McKinsey» в статье [20], где было предложено рассматривать группы специалистов, обладающие эквивалентными функциональными возможностями для решения бизнес-задач, как функциональные дома, а перспективную структуру предприятия — как организованный набор ФД. Учитывая широкое распространение различных форм временного привлечения человеческих ресурсов (например, аутстаффинг² и привлечение фрилансеров³), под сотрудниками $\Phi Д$ будем понимать не только штатных сотрудников фирмы, но также привлекаемых фрилансеров и сотрудников аутстаффинговых фирм.

Отметим, что концепция ФД представляет собой дальнейшее развитие матричной и проектной организационных структур, с одной стороны, и стандартизации рабочих процессов, результатов труда (например, параметров изделия и норм выработки), а также навыков и знаний сотрудников, с другой [21]. Функциональные дома играют роль пулов идентичных ресурсов, из которых составляются временные или постоянные рабочие группы для осуществления бизнес-деятельности в виде выполнения регулярных заданий, проектов, программ и др. Свое целевое назначение — создание результата — сотрудники осуществляют в рамках не ФД, а рабочих групп.

Фактически управление ЧК представляет собой управление бизнес-возможностями функциональных домов — функциональными возможностями специалистов и их численностью. Поэтому аналогично популярным концепциям управления произ-

водственными активами, управления жизненным циклом производственных активов, целесообразно говорить об управлении жизненным циклом функциональных домов и жизненным циклом их сотрудников.

Управление человеческим капиталом будем понимать как определение необходимого набора ФД (требований к функциональным возможностям сотрудников) и численности каждого из них соответственно потребностями бизнеса ПЗЧК в различные моменты времени.

Выбор набора ФД задается характером бизнеса и поэтому вынесен из рассмотрения. В то же время, численный состав ФД (и предприятия в целом) является основным фактором, определяющим операционные затраты и поэтому непосредственно и существенно влияющим на финансовые результаты. Вопрос управления ЧК со стороны менеджмента ПЗЧК распадается на два направления:

- управление жизненным циклом сотрудников в рамках карьеры на предприятии: прием и адаптация, обучение, перемещения и т. д.;
- управление численностью ФД путем найма и содержания штатных сотрудников, а также пользование услугами аутстаффинговых фирм или фрилансеров.

Фактически оптимальным является управление, которое обеспечивает наилучшее соответствие численности и функциональных возможностей сотрудников различных ФД потребностям бизнеса.

Для обеспечения практической применимости в бизнесе модель должна основываться на концепшиях и соответствующих соотношениях управленческого учета, что предполагает достаточно высокую точность и детальность описаний. Однако, бизнес сложен, и моделирование его элементов всегда сопровождается существенными предположениями и упрощениями. Явления и параметры, которыми мы оперируем, относятся к социальной, психологической сфере, физически неизмеримы и точно неизвестны, а параметризация зависимостей часто условна. Возможности применения «статистических усреднений» тоже ограничены, так как практический интерес представляет не массово повторяющиеся явления, а конкретные реализации бизнеса именно в этом периоде времени и при этих внешних условиях.

Поэтому очень важно обеспечение устойчивости выводов к неточно определенным параметрам и закономерностям, используемым в модели, для применимости их на практике. Для этого воспользуемся интервальными оценками вместо точечных, проанализируем чувствительность решений к значениям параметров, преднамеренно «не повышая» подробность и точность моделей.

² Аутстаффинг (англ. out — «вне» + staff — «штат») — способ управления персоналом, предполагающий оказание услуг в форме предоставления в распоряжение заказчика определенного числа работников, не вступающих с ним в какие-либо правовые отношения (гражданско-правовые, трудовые) напрямую, но оказывающих от имени исполнителя определенные услуги (работы) по месту нахождения заказчика.

³ Фрилансер (*англ*. freelancer) — свободный работник, частный специалист.



1. БАЗОВАЯ МОДЕЛЬ ЧЕЛОВЕЧЕСКОГО КАПИТАЛА

Далее предлагается базовая модель управления человеческим капиталом как основным ресурсом ПЗЧК — численностью сотрудников, их квалификацией и мотивацией, затратами на их содержание и развитие с учетом неравномерности бизнеса и потребности бизнеса в человеческом капитале.

Предлагаемая модель связывает перечисленные факторы с экономическим эффектом (создаваемым результатом деятельности) ПЗЧК с точки зрения его менеджмента, и в последующих разделах модель будет использована для определения стратегий и оперативного управления человеческим капиталом через управление численностью сотрудников, привлечение субподрядных фирм и фрилансеров.

Модель рассматриваем в дискретном представлении времени с фиксированной длительностью интервала, что достаточно естественно для бизнеса, где всегда существуют отчетные периоды, в дальнейшем t будет означать номер интервала времени.

1.1. Представление человеческого капитала

Будем описывать ПЗЧК в виде совокупности функциональных домов — пулов сотрудников различных компетенций и уровней квалификации, из которых формируются рабочие группы для выполнения работ в рамках заключаемых контрактов или постоянно предоставляемых сервисов. Сотрудников каждого Φ Д считаем статистически одинаковыми, обозначим $n_i(\theta, t)$ — численность сотрудников i-го Φ Д, имеющих стаж $\theta \ge 0$ работы в ПЗЧК на момент окончания t-го интервала времени:

$$n_{i}(0, t) = n_{i}^{+}(t) + n_{i}^{+BH}(t),$$

$$n_{i}(\theta + 1, t + 1) = n_{i}(\theta, t) - n_{i}^{-BH}(\theta, t) - n_{i}^{YB.\Gamma}(\theta, t) - n_{i}^{YB.\Gamma}(\theta, t),$$
(1)

где $n_i^+(t)$ — число вновь нанятых в течение i-го интервала сотрудников i-го Φ Д, $n_i^{+\mathrm{BH}}(t)$ — число переводимых сотрудников в i-й Φ Д из других Φ Д в рамках внутренних перемещений, $n_i^{-\mathrm{BH}}(\theta, t)$ — число переводимых сотрудников из i-го Φ Д в другие Φ Д в рамках внутренних перемещений, $n_i^{\mathrm{yB.II}}(\theta, t)$ — число увольняемых по инициативе предприятия сотрудников, $n_i^{\mathrm{yB.C}}(\theta, t)$ — число увольняющихся по своей инициативе сотрудников.

Движение персонала (1) отличается неопределенностью: $n_i^{\text{yB.c}}(\theta, t)$ — число увольняющихся по

собственному желанию сотрудников (иногда и $n_i^{+\text{вн}}(t)$ — переводимых) не может полностью определяться руководством фирмы.

1.2. Модель потребности в человеческом капитале со стороны бизнеса

Потребность в человеческом капитале будем представлять в виде потока действий различных типов, которые необходимо осуществить для выполнения обязательств бизнеса (возникающих как следствие продаж и/или постоянно предлагаемых потребителям сервисов различных типов). Примерами таких действий служат операции по подготовке документов, написанию фрагментов программного кода, созданию информационных моделей, изготовлению деталей, сборке агрегатов, обслуживанию клиентов и др. Действия одного типа считаем статистически одинаковыми, обозначим $d_i(\tau, t)$ — число на момент окончания t-го интервала времени действий *і*-го типа, исполняемых в $\Pi 3 \Psi K$ в течение $\tau \ge 0$ интервалов времени после их начала:

$$d_{j}(0, t) = d_{j}^{+}(t),$$

$$d_{j}(\tau + 1, t + 1) = d_{j}(\tau, t) - d_{j}^{3AB}(\tau, t); \qquad (2)$$

где $d_j^+(t)$ — число начатых в течение t-го интервала действий j-го типа, $d_j^{\rm 3aB}(\tau,t)$ — число завершенных в течение t-го интервала действий j-го типа.

Выполнение каждого из k-х действий j-го типа, требует определенного объема ЧК — $m_i(k, \theta, t)$ — численностей сотрудников, соответствующих ФД.

Сумма $\sum_{k=1}^{d_j(\tau, t)} m_i(k, \theta, t)$ по всем действиям формирует потребность ЧК по всем ФД предприятия.

Поток $d_j^+(t)$ начала действий и их завершений $d_j^{\rm 3aB}(\tau,t)$ являются неопределенными и порождают будущую неопределенность и неравномерность потребности в ЧК со стороны бизнеса.

2.3. Представление операционной деятельности и создания результата

Рассмотрим операционную деятельность ПЗЧК и создание результата. Очевидны ограничения численности сотрудников, выделяемых для выполнения действий:

$$\sum_{k=1}^{d_j(\tau,t)} m_i(k, \theta, t) \le n_i(\theta, t) \forall i, j, \theta, \tau, t.$$
 (3)



Для оценивания экономического эффекта ПЗЧК будем, следуя подходам [15], учитывать затраты, связанные с формированием и использованием ЧК (по каждому l-му сотруднику), обозначим их: $c_{i,l}^{\rm 3I}(\theta,t)$ — компенсационный пакет (заработная плата, премии и другие материальные льготы) и налоги на него, $c_{i,l}^{\rm of}(\theta,t)$ — обучение и профессиональную подготовку, $c_{i,l}^{\rm ag}(t)$ — прием на работу и адаптацию, $c_{i,l}^{\rm off}(t)$ — поиск сотрудников и подготовку предложений о работе, $c_{i,l}^{\rm nep}(t)$ — перевод внутри компании и адаптацию на новой позиции, $c_{i,l}^{\rm yB,C}(\theta,t)$ — увольнение по инициативе компании, $c_{i,l}^{\rm yB,C}(\theta,t)$ — увольнение по инициативе сотрудника.

Важный фактор формирования ЧК и управления им заключается в создании и поддержании имиджа ПЗЧК как работодателя, поэтому выделим соответствующие затраты $C^{\rm um}(t)$, также обозначим $C^{\rm пр.ч}(t)$ — прочие затраты ПЗЧК, не связанные с персоналом и потому учитываемые единой суммой, $N_i^{\rm off}(t)$ — число подготовленных и выданных потенциальным сотрудникам предложений о приеме на работу.

Следуя сложившимся в современной экономике подходам [22-24], экономический эффект будем понимать как разность между результатом экономической деятельности и затратами, произведенными для его получения. В интересах решения поставленной задачи результат деятельности каждого из сотрудников будем оценивать произведением его продуктивного времени (в течение которого сотрудник был назначен на выполнение требуемых бизнесом действий) на его производительность $v(i, j, \tau, t)$ (результат, создаваемый каждым сотрудником і-го ФД за единицу времени в рамках выполнения действия j-го типа), а результат деятельности всех ресурсов предприятия — как сумму результатов сотрудников всех ФД. Операция суммирования адекватна данным условиям, так как синергетические эффекты совместной работы многих сотрудников (отражающие сложный, системный характер продукта или сервиса любого предприятия, особенно ПЗЧК) учитываются через функции потребностей $m_i(k, \tau, t)$ и соответствующие назначения. Заметим, что предложенный показатель, естественно, не отражает рыночную ценность или результативность деятельности предприятия с точки зрения потребителей. Однако решаемая задача направлена на управление ресурсами, что не требует рассмотрения рыночной ценности, а предложенный показатель как раз и представляет вклад ресурсов в создание результата.

Без ограничения общности и результат, и затраты могут трактоваться и как признанные по завершению работ, и как начисленные параметры незавершенных работ, и как накопленные суммы платежей, или иные финансово-экономические показатели.

Тогда с точки зрения менеджмента $\Pi 3 \Psi K$ экономический эффект человеческого капитала в целом за интервал времени T определится как

$$V_{0,T} = \sum_{t=1}^{T} \sum_{\tau=0}^{\infty} \sum_{j,i} \sum_{k=1}^{d_{j}(\tau,t)} v(i,j,\tau,t) m_{i}(j,\tau,t) - \sum_{t=1}^{T} \sum_{\theta=0}^{\infty} \sum_{i} \left\{ \sum_{l=1}^{n_{i}(\tau,t)} (c_{i,l}^{3\Pi}(\theta,t) + c_{i,l}^{06}(\theta,t)) + \sum_{i=1}^{n_{i}^{YB.\Pi}(\theta,t)} c_{i,l}^{YB.\Pi}(\theta,t) + \sum_{i=1}^{n_{i}^{YB.C}(\theta,t)} c_{i,l}^{YB.C}(\theta,t) \right\} - \sum_{t=1}^{T} \sum_{i} \left\{ \sum_{l=1}^{n_{i}^{+}(t)} c_{i,l}^{3\Pi}(t) + \sum_{l=1}^{n_{i}^{+BH}(t)} c_{i,l}^{\Pi ep}(t) + \sum_{l=1}^{N_{i}^{off}} c_{i,l}^{off}(t) \right\} - \sum_{t=1}^{T} \left\{ C^{\text{MM}}(t) + C^{\Pi pq}(t) \right\}.$$

$$(4)$$

Соотношения (1)—(4) представляют собой основные соотношения управленческого учета, они точны с финансовой и экономической точки зрения, составлены без упрощающих предположений и пригодны для учета фактических результатов деятельности ПЗЧК. Для анализа и выбора стратегий управления и оперативного управления предприятием необходимо прогнозировать будущие поведение ПЗЧК ($V_{0,T}$) в зависимости от влияющих факторов и оптимизировать экономический эффект.

Часть переменных модели (1)—(4) известны (управляемые или неуправляемые), другая часть — стохастические. Оптимизационная задача относительно экономического эффекта $V_{0,T}$ (4) с учетом условий (1)—(3) может быть сформулирована в виде:

найти значения управляемых переменных $N_i^{o\!f\!f}(t),\ n_i^{-{\scriptscriptstyle {
m BH}}}(\theta,t),...,\ C^{{\scriptscriptstyle {
m IM}}}(t),$ удовлетворяющие ограничениям (1)—(3) и обеспечивающие $\overline{V}_{0,\,T} o {\rm max}.$

Однако данная задача «в точной формулировке» не столько затрудняется размерностью (большое число параметров, меняющих значения во времени), сколько имеет мало бизнес-смысла из-за «истинной неопределенности бизнеса» [25]. «Истинная неопределенность» приводит к тому, что обоснованные предположения о будущем поведении бизнеса не могут быть строгими и подробно



учитывающими «все детали» бизнеса (например, детальную зависимость от времени рынка труда, влияющего на условия найма персонала, уровни заработной платы и интенсивность увольнений по инициативе сотрудников).

Поэтому для решения поставленных задач анализа и выбора стратегий и оперативного управления ЧК примем допущения, обеспечивающие возможность практического использования основных соотношений базовой модели.

2.4. Экономическая модель жизненного цикла сотрудника

Введем допущения и преобразуем базовую модель к виду, удобному для анализа экономики жизненного цикла сотрудников.

Будем рассматривать величины $d_j(\tau, t)$ и $n_j(\theta, t)$ как случайные процессы с дискретным временем.

Переменную $n_i^+(t)$, отражающую число принятых сотрудников, с хорошей точностью можно считать случайной биномиально распределенной величиной с постоянными (предположение о «квазистационарности») параметрами $\langle N_i^{\it off}, \, \pi_i^{\it off}(\Lambda) \rangle$ число сделанных потенциальным сотрудникам предложений о приеме на работу и вероятность принятия предложения соответственно. Численности сотрудников, переводимых внутри ПЗЧК и увольняемых по инициативе компании, $n_i^{+BH}(t)$, $n_i^{-\text{BH}}(t)$ и $n_i^{\text{ув.п}}(\theta, t)$, являются известными детерминированными величинами, с помощью которых управляется ПЗЧК. Число увольняющихся по своей инициативе сотрудников $n_i^{\text{yB.c}}(\theta, t)$ в большинстве случаев с хорошей точностью можно считать случайной величиной, биномиально распределенной с параметрами $\langle (n_i(\theta, t) - n_i^{-BH}(\theta, t)$ $n_i^{ ext{yB.C}}\left(heta,\ t
ight);\ \pi_i^{ ext{yB.C}}\left(\Lambda,\ heta
ight)>,\ ext{где}\ \pi_i^{ ext{yB.C}}\left(\Lambda,\ heta
ight)$ — вероятность увольнения сотрудников по собственной инициативе в течение одного интервала времени.

Важными параметрами динамики численности Φ Д являются вероятности принятия предложения о приеме и увольнения сотрудников по их инициативе в течение одного интервала времени. Для обеих вероятностей $\pi_i^+(\Lambda)$ и $\pi_i^-(\Lambda,\theta)$ в качестве аргумента использован символ Λ , отражающий зависимости вероятностей от того, как каждый из сотрудников оценивает для себя полезность работы в ПЗЧК по сравнению с другими альтернативами рынка труда. Эта полезность, в свою очередь, зависит от множества факторов, таких как размер компенсационного пакета, имидж фирмы и другие, проанализируем эти вероятности ниже.

Затраты всех видов $c_{i,l}^*(\theta, t)$ по отдельным l-м сотрудникам будем представлять независимыми случайно распределенными величинами с известными средними значениями $\bar{c}_i^*(\theta)$ и дисперсиями $\sigma_{*i}(\theta)$, принимая во внимание статистическую эквивалентность сотрудников одного i-го $\Phi Д$.

Результат деятельности сотрудника на каждом интервале определен выше как произведение производительности сотрудника на время, в течение которого сотрудник работал в рамках назначений, отвечающих потребностям бизнеса, — время его продуктивной загрузки. Учитывая статистическую эквивалентность сотрудников, введем среднее значение результата сотрудника $\bar{v}_i(\tau)$ как произведение его средней производительности на среднее время назначений.

С учетом введенных допущений перейдем к среднему значению $V_{0,\,T}$ — эффекту от сотрудников одного ФД (индекс, указывающий номер ФД опустим) в течение интервала времени $[0,\,T]$:

$$\begin{split} V_{0,T} &= \sum_{t} \left\{ \sum_{i,\,\tau}^{n(\tau,\,t)} [\,\bar{v}_i(\tau) - c_{_{3\Pi,i}}(\tau,\,t) - c_{_{\mathrm{TPeH},i}}(\tau,\,t)] \, - \right. \\ &- \sum_{i}^{n_{\mathrm{o}\varphi\varphi}(t)} [\,c_{\mathrm{o}\varphi\varphi,i}(t)] \, - \sum_{i}^{n(0,\,t)} [\,c_{_{\Pi\mathrm{p},i}}(t)] \, - \\ &- \sum_{i,\,\tau}^{n_{_{\mathrm{YB,C}}}(\tau)} [\,c_{_{\mathrm{YB,C},i}}(\tau,\,t)] \, - \sum_{i,\,\tau}^{n_{_{\mathrm{YB,K}}}(\tau)} [\,c_{_{\mathrm{YB,K},i}}(\tau,\,t)] \, - c_{_{\mathrm{COII}}}(t) \, \right\}. \end{split}$$

Усредняя по сотрудникам и по времени, предполагая стационарный режим, получим эффект:

$$V = n_{\text{opp}} \pi^{+}(\Lambda) \left\{ \sum_{\theta=0}^{\infty} \varphi(\theta) S(\theta) - S_{\text{Haŭm}} \right\} - c_{\text{coll}}, \quad (5)$$

где $\varphi(\theta) = \prod_{\omega=0}^{\theta-1} (1-\pi^-(\omega,\Lambda))(1-\delta_{\mathrm{yB.K}}(\omega))$ — доля сотрудников Φ Д, доработавших до стажа θ , по отношению к принятым $n_{\mathrm{o}\varphi\varphi}\pi^+(\Lambda); \ S(\theta) = \bar{\nu}(\theta)$ — $-c_{_{3\Pi}}(\theta) - c_{_{0}\mathrm{буч}}(\theta) - \pi^-(\theta,\Lambda)c_{_{\mathrm{yB.C}}}(\theta) - \delta_{_{\mathrm{yB.K}}}(\theta)c_{_{\mathrm{yB.K}}}(\theta)$ — экономический эффект от сотрудника на θ -м периоде его работы;

 $S_{{
m Haйm}} = c_{{
m o} {
m d} {
m d}}/\pi^+(\Lambda) + c_{{
m np}} - {
m c}$ тоимость приема сотрудника и адаптации его на рабочем месте;

 $c_{\rm cou}$ — затраты на персонал в целом (без отнесения к конкретному сотруднику), в частности на создание имиджа предприятия, социального климата и т. д.

Соотношения (5) пригодны для решения задачи оптимизации экономического эффекта как от $\Phi Д$ в целом $V \rightarrow$ max, так и одного сотрудника на всем



его жизненном цикле $\sum_{\theta=0}^{\infty} \phi(\theta) S(\theta) - S_{\text{найм}} \to \max$, и будут далее учитываться в методике анализа и управления ЧК.

2. МЕТОДИКА АНАЛИЗА И УПРАВЛЕНИЯ ЧЕЛОВЕЧЕСКИМ КАПИТАЛОМ

Сформулируем экономическую методику управления человеческим капиталом, основанную на предложенных математических моделях.

Управление человеческим капиталом осуществляется под влиянием внешних и внутренних по отношению к ПЗЧК бизнес-факторов. Это, прежде всего, размеры компенсационных пакетов сотрудников в фирме и «на рынке», имидж фирмы на рынке труда, перспективы развития сотрудника в процессе работы на фирме, социальные условия работы и многие другие аспекты, которые могут быть агрегированы в понятие полезности, употребленное в виде аргумента Л в выражениях для вероятностей $\pi_{i}^{+}(\Lambda)$ и $\pi_{i}^{-}(\Lambda, \theta)$. Их воздействие «подталкивает» менеджмент находить область оптимального функционирования и поддерживать фирму — параметры ФД — в окрестности оптимума, находясь в определенном равновесии с «рынком». Фактически менеджмент или поддерживает фирму в окрестности текущего равновесия-оптимума, или, если условия (внутренние и/или внешние) изменились, должен как можно быстрее обнаружить изменение условий, спрогнозировать новую область равновесия-оптимума и «сдвинуть» параметры ФД в эту новую область.

Предлагаемая методика помогает обоснованно оценивать и прогнозировать эту зону и реагировать на изменение внешних и внутренних условий, когда они (внешние и внутренние условия) «сдвигают» эту зону.

Таким образом, методика анализа и формирования рекомендаций по управлению ЧК должна состоять из следующих шагов.

Шаг 1. Собираем данные по всем ФД — численности $(n_i(\theta, t))$, производственную загрузку и потребности $(d_j(\tau, t)$ и $m_i(j, \tau, t))$, создаваемый сотрудниками результат $(\bar{v}_i(\tau))$; затраты $(c_i(\tau))$ — виды стоимости персонала для ПЗЧК. Недостающие данные задаем экспертным способом.

Шаг 2. Адаптируем модель и выделяем наиболее влияющие параметры.

Шаг 3. Проверяем близость к оптимуму в рамках текущей стратегии. Верифицируем результат проверки. Определяем наиболее критичные параметры, которые необходимо отслеживать и управлять, чтобы оставаться в окрестности оптимума. При невозможности балансирования в окрестности оптимума (существенное изменение внешних и внутренних условий) необходимо повторить более ранние шаги, начиная с первого.

Шаг 4. Рассматриваем целесообразность смены модели — варианты:

- переход от одной стратегии к другой;
- изменение структуры ФД;
- переход от собственных ресурсов к подряду, привлечение аутстаферов и фрилансеров, определение их долей.

3. ПРИМЕР АНАЛИЗА

Проиллюстрируем применение изложенной методики на примере одного из $\Phi Д$ крупной консалтинговой фирмы.

Шаги 1 и 2. Анализ учетных данных показал, что численность ФД стационарна и колеблется в пределах 40—60 чел. Сотрудники имеют стабильную среднюю загрузку и в течение всего времени работы в компании создают результат, много больший затрат на их содержание. Основная доля затрат на персонал приходится на компенсационный пакет, размер пакета практически не изменяется ни в зависимости от календарного времени, ни от стажа работы в фирме, изменения в пределах 1...3 % не могут считаться значимыми:

$$\begin{split} \bar{v}\left(\theta\right) &\approx \mathrm{const} = v; \quad c_{_{3\Pi}}(\theta) \approx \mathrm{const} = c_{_{3\Pi}.0}; \\ c_{_{0}\Phi\Phi} &= \alpha_{_{0}\Phi\Phi}c_{_{3\Pi}.0}; \quad c_{_{\Pi p}} \approx \alpha_{_{\Pi p}}c_{_{3\Pi}.0}; \\ c_{_{0}\theta y q} &= c_{_{yB.C}} = c_{_{yB.K}} \approx 0; \quad \alpha_{_{\Pi p}} \approx 0,3; \\ \alpha_{_{0}\Phi\Phi} &\approx 0,5. \end{split} \tag{6}$$

Соотношения (6) записаны с учетом всех налогов, пенсионных и других отчислений.

Увольнения по инициативе фирмы пренебрежимо малы, а динамика трафика сотрудников описывается выборочной гистограммой (рис. 1)

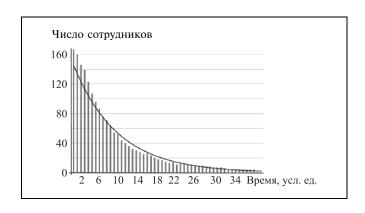


Рис. 1. Гистограмма распределения числа сотрудников по стажу работы



распределения числа сотрудников в зависимости от времени, проработанного ими в фирме (по горизонтали, в условных единицах).

Видно, что время работы сотрудника с хорошей точностью может быть представлено случайной экспоненциально распределенной величиной, т. е. $\pi_i^-(\Lambda, \theta)$ не зависит от стажа θ и $\pi_i^-(\Lambda, \theta)$ = $\pi_i^-(\Lambda) = \pi_0^- \approx 1/10...1/13$. Вероятность принятия будущим сотрудником предложения о работе, оцененная по историческим данным, составляет $\pi_i^+(\Lambda) = \pi_0^+ \approx 0.95$.

Параметр Λ , отражающий оценку каждым сотрудником полезности работы в фирме по сравнению с другими альтернативами рынка труда, в данном случае целесообразно заменить на $c_{_{3\Pi}}$, так этот параметр единственный ненулевой из используемых в модели в данном примере.

Предполагаем, что менеджмент стремится поддерживать постоянную численность Φ Д, тогда баланс трафика сотрудников описывается соотношением $\pi^-(c_{_{3\Pi}})M_{_{\rm CP}}=n_{_{{\rm O}}\varphi\varphi}\pi^+(c_{_{3\Pi}})$, где $M_{_{\rm CP}}$ — средняя численность Φ Д, фактически менеджмент поддерживает баланс трафика сотрудников между фирмой (Φ Д) и рынком.

Шаг 3. С учетом сделанных замечаний и обозначений эффект от ФД получится в форме:

$$\begin{split} V &= \mathit{M}_{\rm cp} \pi^{-}(c_{_{3\Pi}}) \times \\ &\times \left\{ \frac{1}{\pi^{-}(c_{_{3\Pi}})} (v - c_{_{3\Pi}}) - \frac{\alpha_{\rm opp} c_{_{3\Pi}}}{\pi^{+}(c_{_{3\Pi}})} - \alpha_{\rm пp} c_{_{3\Pi}} \right\} - c_{_{\rm COII}}. \end{split}$$

Перепишем это выражение в удобной для анализа форме:

$$V = (M_{\rm cp} v - c_{\rm cou}) -$$

$$- M_{\rm cp} c_{\rm 3\Pi} \left\{ 1 + \alpha_{\rm o\phi\phi} \frac{\pi^{-}(c_{\rm 3\Pi})}{\pi^{+}(c_{\rm 3\Pi})} + \alpha_{\rm \Pi p} \pi^{-}(c_{\rm 3\Pi}) \right\}.$$
 (7)

Выражение (7) отражает все значимые в данном случае параметры и может быть использовано для решения задач практического управления ЧК, а именно, формирования рекомендаций по оптимальному уровню затрат на компенсационный пакет, для чего следует найти значение $c_{3\pi}$, минимизирующее

$$Z(c_{_{3\Pi}}) = c_{_{3\Pi}} \left\{ 1 + \alpha_{_{0} \Phi \Phi} \frac{\pi^{-}(c_{_{3\Pi}})}{\pi^{+}(c_{_{3\Pi}})} + \alpha_{_{\Pi p}} \pi^{-}(c_{_{3\Pi}}) \right\} \rightarrow \min.$$

Решение этой простой с математической точки зрения задачи затрудняется тем, что зависимости обеих вероятностей от затрат на компенсационный пакет $\pi^-(c_{_{3\Pi}})$ и $\pi^+(c_{_{3\Pi}})$ отражают социально-психологические явления и не могут быть достоверно

оценены, измерены или получены из каких-либо объективных законов. Однако соображения здравого смысла позволяют сформулировать ряд их свойств (будем ссылаться на эти свойства как на С1):

— при
$$c \gg c_{3\pi.0}$$
; $\pi^{-}(c) \approx 0$, $\pi^{+}(c) \approx 1$;

— при
$$c \ll c_{3\pi,0}$$
; $\pi^{-}(c) \approx 1$, $\pi^{+}(c) \approx 0$;

$$-\pi^{-}(c_{_{3\Pi}\,0})=\pi_{0}^{-}$$
 и $\pi^{+}(c_{_{3\Pi}\,0})=\pi_{0}^{+}$;

— функции $\pi^{\pm}(c)$ не имеют особых точек и монотонно не убывают/не возрастают по c.

Также менеджмент фирмы, основываясь на опыте, может оценить интервал «разумных» значений $c_{_{3\Pi},0}-\Delta c_1 \leq c_{_{3\Pi}} \leq c_{_{3\Pi}}+\Delta c_2$, выход, за пределы которого не имеет смысла по бизнес-соображениям

Выберем семейство «легко параметризуемых» функций $\pi^{\mp}(c;X)$ и подберем значения параметров X так, чтобы выполнялись требуемые свойства C1. Для вероятностей $\pi^{\mp}(c;X)$ удобно воспользоваться семейством логистических функций $f(z)=(1+\exp(-z))^{-1}$. С учетом условий $\pi^{-}(c_{_{3\Pi}.0};X)=\pi_{0}^{-}$, $\pi^{+}(c_{_{3\Pi}.0};X)=\pi_{0}^{+}$ функции вероятностей выразим:

$$\pi^{-}(c; X) = \frac{\pi_{0}^{-}}{\pi_{0}^{-} + (1 - \pi_{0}^{-}) \exp(x_{1}(c - c_{3\pi, 0}))},$$

$$\pi^{+}(c; X) = \frac{\pi_{0}^{+}}{\pi_{0}^{+} + (1 - \pi_{0}^{+}) \exp(x_{2}(c_{opt,0} - c))},$$

где x_1 и x_2 — положительные параметры.

Теперь оптимизационная задача может быть сформулирована в виде:

$$\begin{cases} Z(c, x_1, x_2) = \\ = c \left\{ 1 + \alpha_{\text{opp}} \frac{\pi^-(c; x_1)}{\pi^+(c; x_2)} + \alpha_{\text{np}} \pi^-(c, x_1) \right\} \rightarrow \min, \\ c_{3\Pi, 0} - \Delta c_1 \le c \le c_{3\Pi, 0} + \Delta c_2, x_1 > 0, x_2 > 0. \end{cases}$$

Полученная задача нелинейного программирования относительно трех переменных (c, x_1, x_2) с гладкой целевой функцией может быть исследована и решена аналитически. Однако для применимости в бизнесе более целесообразно ее численное решение с графической интерпретацией (рис. 2). На всех графиках горизонтальная ось соответствует затратам на компенсационный пакет в относительных единицах $c/c_{3\pi.0}-1$, нулевая точка оси соответствует $c_{3\pi.0}$. Семейства кривых представляют графики зависимостей затрат $Z(c, x_1, x_2)$ и вероят-



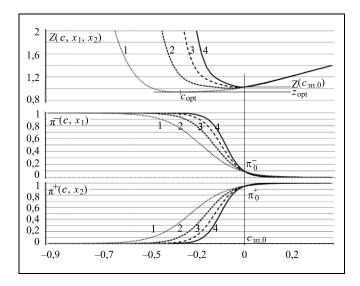


Рис. 2. Графики зависимостей $Z(c, x_1, x_2); \pi^-(c, x_1); \pi^+(c, x_2)$

ностей $\pi^-(c; x_1); \pi^+(c; x_2)$ от c для нескольких значений параметров x_1, x_2 .

Кривые *1* на графиках вероятностей отражают слабо консолидированный рынок труда — т. е. когда диапазон приемлемых для сотрудников размеров компенсационного пакета (зона, в которой вероятность существенно больше 0 и меньше 1, достаточно велика), напротив, кривые *4* отвечают высококонсолидированному рынку труда — узкому разбросу размеров компенсаций. Получив семейства кривых, менеджер на основе своего экспертного мнения определяет, к какому виду рынка относится текущая ситуация.

Если рынок консолидирован (кривые 4), то потенциальное уменьшение затрат $Z(c, x_1, x_2)$ очень незначительно: минимум кривой $4\,Z(c, x_1, x_2)$ практически находится в точке $c_{_{3\Pi,0}}$. Очевидно, в этом случае $\Phi Д$ находится в оптимальном равновесии с рынком и изменения в системе нецелесообразны.

Для противоположного случая (кривые 1) $Z_{\rm opt}$ может отличаться от $Z(c_{\rm 3II.0})$ на существенную величину, а $c_{\rm opt}$ — от $c_{\rm 3II.0}$ (в нашем примере на 20~%). В этом случае, вероятно, менеджмент примет решение проверить возможность перехода к точке оптимума, изменив размер компенсационного пакета.

Наиболее значимым параметром для контроля в данном случае является размер компенсационного пакета «на рынке», изменение которого может существенно нарушить баланс трафика. Поэтому этот параметр требует не только отслеживания, но и прогнозирования.

Изложенный пример анализа может был полезен как для проверки достижения оптимальной области равновесия, так и для прогнозирования таких областей в случае изменения внешних или внутренних условий.

Шаг 4. После того, как возможности оптимизации путем управления параметрами затрат исчерпаны, рассмотрим способ комбинирования штатных сотрудников с привлечением фрилансеров и аутстаферов для обеспечения неравномерной потребности в ресурсах со стороны бизнеса.

На рис. 3 представлена выборочная траектория потребности в сотрудниках ФД (вертикальная ось) в зависимости от времени (горизонтальная ось).

Рассмотрим случай, когда менеджмент поддерживает некоторую постоянную M_{Φ} численность Φ Д за счет штатных сотрудников, а превышающие это уровень потребности покрывает привлечением внешних ресурсов (фрилансеров и аутстаферов). Достоинство таких внешних ресурсов состоит в возможности мгновенного привлечения и отказа от продолжения работ без дополнительных затрат, недостаток — в более высокой текущей цене ресурса.

Смысл реализации такой стратегии состоит в возможности сэкономить на оплате временно не-используемых ресурсов при неравномерной потребности в них.

Неравномерность потребности с течением времени будем моделировать последовательностью независимых одинаково распределенных величин ξ_I (потребности в сотрудниках) с функцией распределения вероятности $P(\xi)$. В данной работе не рассматривается вопрос прогнозирования потребности, поэтому такая простая модель вполне достаточна. Очень важно, что из-за простоты модели решения и рекомендации, получаемые с ее помощью, устойчивы к неточностям в предположениях и задаваемых параметрах.



Рис. 3. Выборочная потребность в сотрудниках $\Phi Д$ в зависимости от времени



В этом случае эффект от ФД определится как

$$V = (M_{\rm cp} v - c_{\rm cou}) - M_{\rm \phi} \hat{c}_{\rm mt} -$$

$$- c_{\rm out} \sum_{\xi = M_{\rm \phi}}^{\infty} (\xi - M_{\rm \phi}) P(\xi),$$
(8)

где $c_{\rm out}$ — затраты на одного внешнего специалиста в течение одного интервала времени, $M_{\rm cp}$ =

$$=\sum_{\xi=0}^{\infty}\xi N(\xi)$$
 — среднее значение потребности,

$$\hat{c}_{\text{шт}} = c_{\text{зп}} \left(1 + \alpha_{\text{офф}} \frac{\pi^{-}(c_{\text{зп}})}{\pi^{+}(c_{\text{зп}})} + \alpha_{\text{пр}} \pi^{-}(c_{\text{зп}}) \right) -$$
затраты

на одного штатного сотрудника.

Задача оптимизации (8) сводится к минимизации затратной части:

$$Z(M_{\phi}) =$$

$$= \alpha_{\text{out}} M_{\phi} + \sum_{\xi = M_{\phi}}^{\infty} (\xi - M_{\phi}) P(\xi) \to \text{min}; \qquad (9)$$

где $\alpha_{\text{out}} = \hat{c}_{\text{шт}}/c_{\text{out}}$.

Она легко решается численно, так как функция $\sum_{\xi=M_{\Phi}}^{\infty} (\xi-M_{\Phi}) P(\xi)$ монотонно спадает с рос-

том M_{Φ} , и оптимум легко находится перебором значений M_{Φ} .

На основе исторических данных (см. рис. 3) была построена выборочная функция распределения вероятностей $P(\xi)$, которая использовалась для решения задачи (9). На рис. 4 представлены графики зависимостей $Z(M_{\phi})$ (вверху) для различных значений $\alpha_{\text{оит}}$ и выборочная функция $P(\xi)$ (внизу).

При низкой стоимости внешних специалистов по сравнению со штатными ($\alpha_{\rm out} \cong 1; \ \hat{c}_{\rm шT} = c_{\rm out}$) функция затрат $Z(M_{\phi})$ не имеет минимума: принимает постоянное значение, равное среднему значению потребности.

При высокой стоимости внешних специалистов ($\alpha_{\rm out}=0.1\ll1$; $c_{\rm out}=10\,\hat{c}_{\rm шT}$) точка минимума лежит ближе к верхней границе диапазона возможных значений потребности, что логично, и минимум слабо выражен.

При наличии исторических данных и справедливости предположения о сохранении в будущем существующих трендов применения выборочной функции распределения вероятностей является наиболее адекватным и простым подходом для получения оптимальных значений численности штатных сотрудников $M_{\rm th}$.

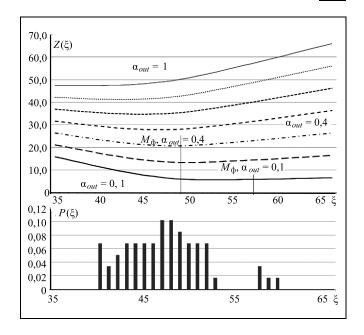


Рис. 4. Зависимости затрат $Z(M_{\Phi})$ и выборочная функция распределения $P(\xi)$

При отсутствии таких данных, например, в случаях создания новых бизнесов и Φ Д или при ожидании значительных изменений внешних или внутренних факторов, функция $P(\xi)$ может быть аппроксимирована распространенными распределениями, что позволит получить предварительные оценки значений M_{Φ} .

Численное моделирование показало, что в данном примере замена выборочного распределения на биномиальное с эквивалентными двумя моментами приводит к отличию получаемых значений M_{ϕ} для двух распределений не более чем на единицу. Это позволяет использовать биномиальное распределение для получения предварительных оценок численности штатных сотрудников при отсутствии исторических данных.

ЗАКЛЮЧЕНИЕ

Предложенная методика позволяет формировать практические рекомендации по управлению человеческим капиталом через управление отдельными видами затрат на персонал и численностью ФД предприятия, создающего и интенсивно использующего знания.

Важно, что аналитические выражения в моделях получены из «точных» соотношений управленческого учета и реального бизнеса путем внесения ограниченного набора предположений.

Разработанная модель учитывает основные драйверы неопределенности функционирования



предприятия, интенсивно использующего знания и человеческий капитал, — неравномерность потребности и трафик сотрудников.

Тема управления человеческим капиталом весьма актуальна, поэтому предложенный подход целесообразно развивать в различных направлениях, взяв в качестве основы базовую модель (1)—(4).

Для бизнес-применений интересны такие направления развития методики, как, например, исследование изменения эффективности сотрудников в течение жизненного цикла в компании, учет различного уровня мотивации сотрудников, разработка методов управления трафиком сотрудников через принудительное увольнение в соответствии с различными критериями, ограничение численности ФД ниже среднего уровня потребности на уровне гарантированной загрузки и сознательный отказ от реализации части бизнеса и другие сценарии управления человеческим капиталом.

Отдельного исследования также заслуживают вопросы аппроксимации неопределенного трафика сотрудников и неравномерной потребности удобными для анализа моделями случайных процессов. Такие модели должны позволять адекватно прогнозировать бизнес-явления и рассматривать в том числе нестационарные процессы, отвечающие случаям трансформаций предприятий.

ЛИТЕРАТУРА

- Human capital management definition. URL: http://blogs.gartner.com/it-glossary/hcm-human-capital-management / (дата обращения: 20.03.2016).
- Smith D., Silverstone Y. and Lajtha A. A new lens on business advantage: Human capital strategy and the drive for high performance / Accenture, 2013.
- 3. *Palaciosy M.* Human Capital as an Asset Class Implications From a General Equilibrium Model. Owen Vanderbilt University. Graduate School of Management. URL: http://www2.owen.vanderbilt.edu/miguel.palacios/index_files/Palacios_Human_Capital_as_an_Asset_Class.pdf (дата обращения: 12.02.2016).
- Blackburn K. and Varvarigos D. Human Capital Accumulation in a Stochastic Environment: Some New Results on the Relationship Between Growth and Volatility. — URL: http:// www.socialsciences.manchester.ac.uk/medialibrary/cgbcr/discussionpapers/dpcgbcr74.pdf (дата обращения: 15.02.2016).
- 5. Fan X., Seshadri A., Taber C. Estimation of a Life-Cycle Model with Human Capital, Labor Supply and Retirement. University of Chicago. Department of Economics. Workshops. URL: http://www.ssc.wisc.edu/~aseshadr/WorkingPapers/FST.pdf (дата обращения: 12.09.2016).
- Wu Y. Organizational Structure and Product Choice in Knowledge-Intensive Firms // Management Science. 2015. Vol. 61, N 8. P. 1830—1848.

- Mincer J. Investment in Human Capital and Personal Income Distribution // Journal of Political Economy. — 1958. — Vol. 66, N 4. — P. 281—302.
- URL: http://sebokwiki.org/wiki/Enterprise_(glossary) (дата обращения: 12.09.2016).
- 9. *Белов М.В.* Системно-инженерные и экономические аспекты управления жизненным циклом // Управление большими системами. 2015. Вып. 56. С. 6—65.
- Йон X. Вовлеченность персонала в России. URL: http://www.klex.ru/glg 11-02-2016 (дата обращения: 17.01.2016).
- Kahn W.A. Psychological Conditions of Personal Engagement and Disengagement at Work // Academy of Management Journal. — 1990. — Vol. 33, N 4. P. 692—724.
- 12. *Исследование* вовлеченности. URL: http://axesgroup.ru/engagement/ (дата обращения: 11.02.2016).
- Смирнов В.Т., Сошников И.В., Романчин В.И., Скоблякова И.В. Человеческий капитал: содержание и виды, оценка и стимулирование: монография / под ред. В.Т. Смирнова. — М.: Машиностроение-1; Орел: ОрелГТУ, 2005. — 513 с.
- 14. *Thie H.A.*, *Browh R.A.* Future Career Management Systems for U.S. Military Officers. RAND Corp, 1994. 410 р. URL: https://www.rand.org/content/dam/rand/pubs/monograph_reports/2007/MR470.pdf (дата обращения: 29.08.2016).
- Flamholtz E. Human resource accounting: advances in concepts, methods, and applications. N.-Y.: Springer Seience + Business Media, 1999. 390 p.
- Flamholtz E., Bullen M., Hua W. Human resource accounting: a historical perspective and future implications // Management Decision. — 2002. — Vol. 40, iss. 10. — P. 947—954.
- Maister D.H. Managing the Professional Services Firm. N.-Y.: Free Press, 1993.
- Becker G.S. Investment in Human Capital: A Theoretical Analysis // Journal of Political Economy. Part 2. 1962. Vol. 70, N 5, 1962. P. 9—49.
- Becker G.S. Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education / Third Edition. — Chicago and London: The University of Chicago Press, 1993.
- 20. De Smet A., Lund S., and Schaninger W. Organizing for the future. Platform-based talent markets help put the emphasis in human-capital management back where it belongs on humans. January 2016. URL: http://www.mckinsey.com/insights/ organization/organizing_for_the_future (дата обращения 29.08.2016).
- 21. *Минцберг Г.* Структура в кулаке: создание эффективной организации / Пер. с англ. под ред. Ю.Н. Каптуревского. СПб.: Питер, 2004.
- Борисов А.Б. Большой экономический словарь / А.Б. Борисов. М.: Книжный мир, 2003. 895 с.
- 23. Лопатников Л.И. Экономико-математический словарь: Словарь современной экономической науки / 5-е изд., перераб. и доп. М.: Дело, 2003. 520 с.
- Knight F.H. Risk, uncertainty and profit. Boston: MA: Hart, Schaffner & Marx; Houghton Mifflin Co. 1921, Reprints of economic classics. — N.-Y.: Sentry Press, 1964. — 381 p.

34



УДК 330

РЕШЕНИЕ ОБРАТНЫХ ЗАДАЧ ЭКОНОМИКИ С ПОМОЩЬЮ МОДИФИЦИРОВАННОГО МЕТОДА ОБРАТНЫХ ВЫЧИСЛЕНИЙ

Е.Б. Грибанова

Отмечено, что обратные задачи возникают при необходимости нахождения оптимальных путей решения вопросов компаний и отвечают на вопрос «как сделать так, чтобы?». Предложены варианты решения обратных задач экономики с помощью модифицированного метода обратных вычислений. Рассмотрены различные виды ограничений на аргументы функций. Показана возможность применения модифицированного метода обратных вычислений для оптимизации функции двух аргументов на заданном интервале. Указано, что полученные результаты могут быть полезны в системах поддержки принятия управленческих решений.

Ключевые слова: обратные вычисления, обратная задача, ограничения, оптимизация, управленческие решения.

ВВЕДЕНИЕ

Задачи, решаемые в области экономики, можно разделить по направлению причинно-следственной связи на прямые и обратные. Прямая задача заключается в установлении следствия по известным причинам, т. е. расчете результата по имеющимся значениям формирующих его величин и виду зависимости, что позволяет оценить текущее состояние системы, сделать прогноз на будущие периоды, исследовать влияние факторов на выходную величину. Примером решения прямой задачи может служить определение выручки предприятия по заданным значениям цены и количества проданного товара.

Обратная задача заключается в установлении причин, приводящих к интересующему следствию, т. е. таком подборе исходных величин, который обеспечил бы заданное значение результата. Полученная информация может быть использована для формирования управленческих решений [1—4]. В качестве примера можно привести определение количества проданного товара и цены, обеспечивающих необходимый прирост выручки.

Решение обратных задач с помощью обратных вычислений — это получение значений приростов аргументов функции на основании задаваемого значения приращения функции и дополнительной информации, поступающей от лица, формирующего решение: коэффициентов относительной важности аргументов, индивидуального коэффициента прироста аргументов. Несмотря на то, что аппарат обратных вычислений был разработан от-

носительно недавно, он уже нашел практическое применение в различных областях, в частности, образовании [5] и экономике [6]. Решение задач подобного рода актуально, поскольку позволяет ответить на вопрос «как сделать так, чтобы?», определить управляющие воздействия для достижения желаемого состояния объекта экономики, что является неотъемлемой функцией систем поддержки принятия решений. Таким образом может быть решена важнейшая задача синтеза целевого управления с системой сбалансированных показателей, где в корне дерева целей расположена стратегическая цель, а на терминальных вершинах оперативные показатели [7]. Под деревом цели понимается иерархическая структура, в которой выделена главная цель — корень дерева и подчиняющиеся ей подцели первого, второго и последующего подуровней, представленные в виде листьев дерева.

Модифицированный метод обратных вычислений заключается в определении аргументов функции на основании ее указанного значения и коэффициентов относительной важности [8] (именно такого рода задачи будут рассмотрены в данной публикации и называться «обратными»). Он предполагает построение уравнения связи между аргументами вида $y = a \pm bx$ и подстановку полученного уравнения в исходное соотношение. Для создания уравнения связи используется минимаксный метод. Суть его заключается в построении уравнения диагонали прямоугольника, образованного минимальными и максимальными значениями величин (рис. 1), при этом в качестве углового коэффициента используется отношение ин-



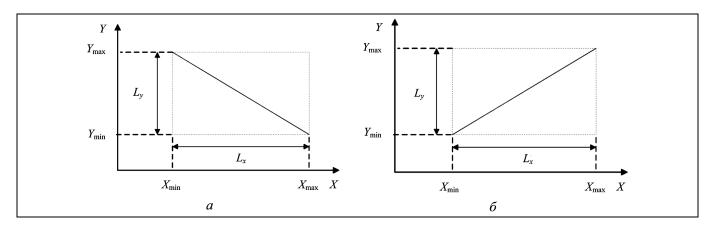


Рис. 1. Зависимость: a — обратная; δ — прямая

тервалов. Так для построения функции обратной зависимости (рис. $1,\ a)$ используются формулы: $b=L_y/L_x,\ a=Y_{\min}+bX_{\min}.$ В случае прямой зависимости (рис. $1,\ \delta$): $b=L_y/L_x,\ a=Y_{\min}-bX_{\min}.$ В модифицированном методе обратных вычислений используются отношение коэффициентов относительной важности в качестве углового коэффициента и исходные данные вместо минимальных значений. В отличие от классического метода обратных вычислений он проще в компьютерной реализации, так как позволяет избежать проверок согласованности дополнительной информации, поступающей от человека: соответствия поставленной цели коэффициентам важности.

Многие задачи экономики предполагают наличие ограничений на значение одного из показателей. Так, например, объем выпуска может быть ограничен производственными ресурсами предприятия. В этом случае дефицит/излишек одного ресурса компенсируется другим. В работе [1] для учета ограничений при использовании обратных вычислений предполагается корректировка коэффициентов относительной важности, в статье [9] описана итерационная процедура оптимизации, которая заключается в последовательном изменении функционала и определении приростов аргументов, что позволяет получить результат с учетом заданных ограничений.

В настоящей работе представлены варианты решения задач экономики с учетом ограничений с помощью модифицированного метода обратных вычислений. В рассмотренных далее задачах предполагается взаимозаменяемость ресурсов.

1. РЕШЕНИЕ ЗАДАЧ ЭКОНОМИКИ С УЧЕТОМ ОГРАНИЧЕНИЙ

Для оценки состояния объекта экономики используются различные показатели, которые могут быть связаны между собой аддитивной мультипликативной и кратной зависимостями [10].

Рассмотрим случай мультипликативной зависимости для функции двух аргументов: выручка равна (рис. 2) r = pc, где p и c — цена и количество товара.

Исходные данные: $r_0=50$ усл. ден. ед., $p_0=10$ усл. ден. ед., $c_0=5$ усл. ед. Необходимо определить значения цены и количества, которые обеспечат размер выручки, равный 100. Установим значения коэффициентов важности приращений аргументов функции: $\alpha=0,4$ и $\beta=0,6$. Полученное с помощью модифицированного метода обратных вычислений решение (рис. 3): b=0,4/0,6=0,667, $a=10-0,667\cdot 5=6,667,$ $(6,668+0,667c_1)c_1=100,$ $c_1=8,228,$ $p_1=12,152.$

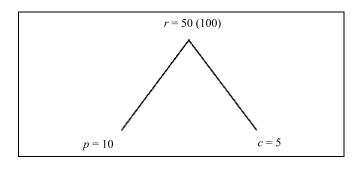


Рис. 2. Формирование выручки

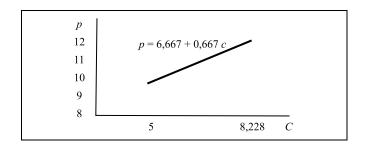


Рис. 3. Зависимость между ценой и количеством



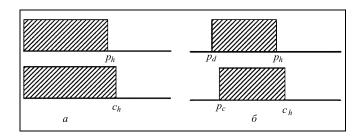


Рис. 4. Виды границ допустимых значений: a — верхняя граница; δ — верхняя и нижняя граница

При нахождении решения не учитывались ограничения, которые могут быть установлены на значения показателей. Можно выделить варианты:

- ограничено значение только одного показателя;
 - ограничены значения двух показателей;
- задана одна граница (верхняя или нижняя, рис. 4, a) допустимых значений;
- установлены две границы (верхняя и нижняя, рис. 4, δ), т. е. указан интервал допустимых значений.

В простейшем случае наличия ограничения значения одного показателя возможна корректировка излишка или дефицита за счет другой величины. В этом случае нахождение решения сводится к следующим шагам.

Шаг 1. Определить исходные данные: c_0 , p_0 α , β , r_1 ; границы: c_h , c_d или p_h , p_d .

Шаг 2. Рассчитать значения искомых величин c_1 , p_1 , обеспечивающих заданное значение функции r_1 с помощью модифицированного метода обратных вычислений. Если полученные значения удовлетворяют ограничениям, то они являются решением задачи, иначе происходит переход к следующему шагу.

Шаг 3. Присвоить показателю граничное значение. Если для величины установлена и верхняя и нижняя границы, то используется ближайшее к полученному решению значение.

Шаг 4. Подставить значение в исходную функцию и определить значение второго показателя путем решения полученного уравнения.

Рассмотрим выполнение этих шагов на примере. Исходные данные: $r_1=100$ усл. ден. ед., $p_0=10$ усл. ден. ед., $c_0=5$ усл. ед., $\alpha=0,4$ и $\beta=0,6,$ $c_h=8$, т. е. существует ограничение: $c\leq 8$.

Полученное с помощью модифицированного метода обратных вычислений решение ($c_1 = 8,228$, $p_1 = 12,152$) не удовлетворяет ограничению $c \le 8$ (см. рис. 4), так как наблюдается недостаток количества на 0,228 усл. ед., который нужно восполнить за счет цены. Присвоим величине c_1 значение верхней границы и подставим в исходное уравне-

ние: $p_1 \cdot 8 = 100$, т.е. $p_1 = 12,5$. Полученное решение: $c_1 = 8$, $p_1 = 12,5$.

В случае, если обе величины, формирующие результирующий показатель, ограничены, задача усложняется. Рассмотрим вариант, когда для каждого показателя существует только одно ограничение: установлена верхняя или нижняя граница. Первоначально необходимо выполнить проверку на существование решения в заданных интервалах. Для этого нужно проанализировать варианты достижения цели (табл. 1) в соответствии с видом зависимости и направлением изменения величин. Так, например, при необходимости обеспечения положительного прироста результата (т. е. $\Delta C > 0$) путем изменения значений $A(\alpha)$ и $B(\alpha)$ в случае мультипликативной зависимости $C = A(\alpha)B(\beta)$ могут быть рассмотрены два варианта: увеличение значений $A(\alpha)$ и $B(\alpha)$; увеличение аргумента с наибольшим значением коэффициента относительной важности при уменьшении значения переменной с наименьшим коэффициентом важности.

Для мультипликативной модели в случае наличия нижних границ $c \ge c_d$, $p \ge p_d$ и выполнения условия $r(c_d, p_d) > r_1$ решение не может быть найдено в заданных интервалах. Решение будет отсутствовать и при существовании верхних границ значений $c \le c_h$, $p \le p_h$ и выполнении условия $r(c_h, p_h) < r_1$. В этом случае в качестве решения могут быть взяты значения границ: c_h , p_h или c_d , p_d .

При существовании решения в заданных интервалах процедура решения задачи включает в себя следующие шаги.

Шаг 1. Определить исходные данные: α, β, r_1 ; границы: c_h или c_d , p_h или p_d .

Шаг 2. В качестве начальных взять граничные значения: $c_0 = c_d$ или $c_0 = c_h$; $p_0 = p_d$ или $p_0 = p_h$.

Шаг 3. Найти решение с помощью модифицированного метода обратных вычислений. Направление изменения величин обусловливает вид зависимости в уравнении связи: прямая или обратная.

Таблица 1 Варианты достижения цели

Вид зависи-	Усло-	Прирост результата							
мости	вие	•	+	-					
Мультипликативная $A(\alpha)B(\alpha)$	$\alpha > \beta$ $\alpha < \beta$	A^+, B^+ A^+, B^+	$\begin{bmatrix} A^+, B^- \\ A^-, B^+ \end{bmatrix}$	$A^-, B^ A^-, B^-$	A^-, B^+ A^+, B^-				
Аддитивная $A(\alpha) + B(\alpha)$	$\alpha > \beta$ $\alpha < \beta$	A^+, B^+ A^+, B^+	$\begin{bmatrix} A^+, B^- \\ A^-, B^+ \end{bmatrix}$	$A^-, B^ A^-, B^-$	A^-, B^+ A^+, B^-				
Кратная $A(\alpha)/B(\alpha)$	$\alpha > \beta$ $\alpha < \beta$	A^+, B^+ A^-, B^-	$A^+, B^ A^+, B^-$	$A^-, B^ A^+, B^+$	A^-, B^+ A^-, B^+				



При обратной зависимости между аргументами уравнение может не иметь решения или полученные значения могут выходить за границы допустимой области. В этом случае корректируются коэффициенты пропорциональности и данный шаг повторяется.

Так, допустим, что в примере формирования выручки существует ограничение: $c \ge 9, \ p \ge 10,$ исходные данные: $r_1 = 100$ усл. ден. ед., $\alpha = 0,4$ и $\beta = 0,6, \ c_d = 9, \ p_d = 10.$

Присвоим начальным значениям количества и цены товара границы допустимых интервалов: $c_0=9,\,p_0=10.$ Поскольку изменение величин происходит в одном направлении, т. е. для двух аргументов установлена нижняя граница, то используется прямая зависимость между показателями. Полученное с помощью модифицированного метода обратных вычислений решение: $c_1=9,61,\,p_1=10,406.$

Аналогично осуществляется поиск решения, если существует верхняя граница допустимых значений.

В случае, если для одной из величин установлена верхняя граница, а для другой — нижняя, то изменение значений для мультипликативной связи должно происходить в различных направлениях, и при нахождении решения нужно рассмотреть обратную зависимость между аргументами.

Пусть для приведенных ранее исходных данных $c_d = 9$, $p_h = 10$, т. е. существует ограничение: $c \le 9$, $p \le 10$.

При рассмотрении обратной зависимости между аргументами квадратное уравнение корней не имеет, следовательно, решение не может быть получено.

Примем шаг изменения коэффициентов, равным 0,1: $\Delta=0,1$. Тогда новые значения коэффициентов относительной важности $\alpha=0,6+0,1=0,7$ и $\beta=0,4-0,1=0,3$. Полученные значения величин ($c_1=10,873,\ p_1=9,197$) с помощью модифицированного метода обратных вычислений соответствуют ограничениям, следовательно, они являются решением задачи.

Наконец, рассмотрим вариант ограничений в виде допустимых интервалов значений: $p_d \le p \le p_h$, $c_d \le c \le c_h$. Для нахождения решения необходимо воспользоваться итерационной процедурой, состоящей из следующих шагов.

Шаг 1. Определить исходные данные: r_1 , c_h , c_d , p_h , p_d .

Шаг 2. Проверить возможность нахождения решения в заданном интервале.

Шаг 3. Рассчитать интервалы допустимых значений:

$$L_c = c_h - c_d, \quad L_p = p_h - p_d.$$
 (1)

Шаг 4. Определить коэффициенты линейного уравнения:

$$b = L_p/L_c, \quad a = p_d - bc_d. \tag{2}$$

Шаг 5. Подставить уравнение в исходную функцию и найти искомые величины c_1 и p_1 .

После определения значений показателей, обеспечивающих заданный результат r_1 , коэффициенты пропорциональности могут быть вычислены по формулам:

$$\alpha^* = \frac{c_1 - c_0}{(c_1 - c_0) + (p_1 - p_0)},$$

$$\beta^* = \frac{p_1 - p_0}{(c_1 - c_0) + (p_1 - p_0)}.$$
(3)

Рассмотрим решение задачи в случае существования подобных ограничений. Исходные данные: $r_1=100,\ c_h=8,\ c_d=5,\ p_h=15,\ p_d=12,2,\ \text{т. e.}$ 5 $\leq c_1 \leq 8,\ 12,2 \leq p_1 \leq 15.$

Поскольку для рассматриваемой мультипликативной модели выполняется условие $r(c_d, p_d) \le c_1 \le r(c_h, p_h)$, решение может быть найдено в заданных границах.

Расчет интервалов (1): $L_c = 8 - 5 = 3$, $L_p = 15 - 12,2 = 2,8$.

Нахождение коэффициентов линейного уравнения (2):

$$b = L_p/L_c = 2,8/3 = 0,933,$$

$$a = p_d - bc_d = 12,2 - 50,933 = 7,53.$$

Полученное линейное уравнение связи: p = 7,53 + 0,933c. Его подстановка в исходную функцию и нахождение решения: $(7,53 + 0,933c_1)c_1 = 100$. Получим: $c_1 = 7,074$, $p_1 = 14,136$.

Таким образом, значения коэффициентов (3) относительной важности: $\alpha^* = 0,334$, $\beta^* = 0,66$.

Если решение необходимо получить с учетом коэффициентов приоритетности α и β , то необходимо использовать итерационную процедуру, включающую в себя последовательную корректировку коэффициентов.

Шаг 1. Определить исходные значения α и β.

Шаг 2. Найти решение задачи с помощью модифицированного метода обратных вычислений. Если решение получено и соответствует ограничениям, то считать полученные значения решением задачи, иначе — перейти к следующему шагу.

Шаг 3. Рассчитать значения α^* и β^* с помощью итерационной процедуры нахождения решения при наличии интервалов допустимых величин; приросты

$$\Delta_{\alpha} = \frac{\alpha - \alpha^*}{n}, \quad \Delta_{\beta} = \frac{\beta - \beta^*}{n},$$
(4)

где n — число точек разбиения интервала.



Шаг 4. Изменить значения коэффициентов пропорциональности на размер прироста: $\alpha = \alpha + \Delta_{\alpha}$ и $\beta = \beta + \Delta_{\beta}$.

Шаг 5. Найти решение задачи с помощью модифицированного метода обратных вычислений. Если решение получено и соответствует ограничениям, то полученные значения считать решением задачи, иначе — перейти к шагу 4.

Исходные данные: $\alpha=0.23,\ \beta=0.77,\ r_1=100,\ c_h=8,\ c_d=5,\ p_h=15,\ p_d=12.2,\ \text{т. e.}$ существуют ограничения: $5\leq c_1\leq 8;\ 12.2\leq p_1\leq 15.$ Полученное с помощью модифицированного метода обратных вычислений решение: $c=6.56,\ p=15.235.$ Цена не удовлетворяет ограничению $12.2\leq p_1\leq 15,$ поэтому необходимо изменить значения коэффициентов относительной важности.

Пусть n = 3, тогда значения прироста (4)

$$\Delta_{\alpha} = \frac{\alpha^* - \alpha}{n} = \frac{0,334 - 0,23}{3} = 0,035,$$

$$\Delta_{\beta} = -\Delta_{\alpha} = -0,035.$$

Скорректированные значения коэффициентов относительной важности: $\alpha = 0.23 + 0.035 = 0.265$; $\beta = 0.77 - 0.035 = 0.735$.

Полученные значения: $c_1=6,74,\ p_1=14,836$ удовлетворяют ограничениям $5\leq c_1\leq 8,\ 12,2\leq 2$ $\leq p_1\leq 15$ и являются решением задачи.

2. РЕШЕНИЕ ЗАДАЧ ОПТИМИЗАЦИИ

Модифицированный метод обратных вычислений также может быть применен для решения задач оптимизации функций, когда на заданном интервале нужно найти значения аргументов, обеспечивающих минимальное или максимальное значение функции, в частности, для определения начальной точки при многомерной оптимизации. Простейший способ решения такой задачи заключается в применении метода Монте-Карло, когда начальная точка выбирается случайным образом, а также существует подход, основанный на использовании регулярной сетки [11].

Алгоритм метода состоит в следующем.

Шаг 1. Определение исходных данных: функция z(x, y), интервалы допустимых значений, на которых осуществляется поиск минимума: $x_d \le x \le x_h$, $y_d \le y \le y_h$.

Шаг 2. Построение линейного уравнения зависимости аргументов и его подстановка в исходную функцию.

Шаг 3. Минимизация/максимизация полученной функции с помощью одного из методов одномерной оптимизации (равномерного поиска, дихотомии, золотого сечения и т. д.). Подстановка полученной величины в уравнение связи. Полученные значения x^* , y^* могут быть использованы в

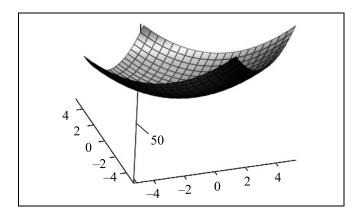


Рис. 5. График функции $z(x, y) = (x - 5)^2 + (y - 1)^2$

качестве начальной точки при применении известных методов оптимизации.

Существующие методы многомерной оптимизации делятся на методы прямого поиска (Хука—Дживса, симплексный, Гаусса—Зейделя и др.) и градиентные (Коши, Флетчера—Ривза и др.) [12]. Основной недостаток градиентных методов состоит в необходимости вычисления производной, а методов прямого поиска — в проведении большого числа итераций для нахождения решения.

Рассмотрим применение данной процедуры для нахождения максимального значения выручки на заданных интервалах. Исходные данные: $r_1=100$, $\epsilon=0,0001$, $\epsilon_{\Delta}=0,04$, $c_h=8$, $c_d=5$, $p_h=15$, $p_d=12,2$. Параметры уравнения: $b=L_p/L_c=2,8/3=0,933$;

 $a = p_d - bc_d = 12,2 - 5 \cdot 0,933 = 7,53.$

Полученное линейное уравнение связи имеет вид: p = 7.53 + 0.933c.

С помощью одномерной оптимизации определим максимум функции: $r(c) = (7,53+0,933\ c)\ c$ на интервале $5 \le c_1 \le 8$. Полученное решение: $c_1 = 8$, $p_1 = 7,53+0,933\cdot 8 = 15$.

Применяя метод Хука—Дживса с полученной начальной точкой, получаем, что решением задачи будет $c_1=8,\ p=15.$

Рассмотрим также применение алгоритма для определения минимума функции $z(x, y) = (x-5)^2 + (y-1)^2$ (рис. 5). Решение задачи оптимизации выполнялось с помощью метода Хука—Дживса (размер шага и погрешность равны 2 и 0,01 соответственно). Для определения начальной точки применялись метод Монте-Карло и модифицированный метод обратных вычислений. При применении метода Монте-Карло моделирование осуществлялось в течение 10 000 реализаций (программа написана на языке VBA). Результаты представлены в табл. 2, где x_d , x_h , x_d и y_d , y_n — границы интервалов первого и второго аргумента, на которых осуществляется поиск минимума; x^* , y^* — ре-

Таблица 2

Сравнение методов поиска начальной точки

Вариант	x_d, x_h	y_d, y_h	<i>X</i> *	<i>y</i> *	I_o	I_m	$I_{m\text{max}}$	$I_{m \min}$	P
1	$ \begin{array}{c c} -2, 10 \\ -10, 10 \\ -25, 25 \end{array} $	-5, 8	4,21	1,73	18	17,17	24	10	0,24
2		-15, 8	4,38	1,53	15	18,63	29	10	0,88
3		-30, 25	4	1,9	12	21,86	36	10	0,99

шение, полученное согласно алгоритму нахождения начальной точки модифицированным методом обратных вычислений.

Так, например, в первом варианте $x_d = -2$, $x_h = 10$, $y_d = -5$, $y_h = 8$. Коэффициенты уравнения: b = (10+2)/(8+5) = 0.923, $a = -2+5\cdot0.923 = 2.615$. Далее осуществляется поиск минимума функции $f(y) = (2.615+0.923y-5)^2+(y-1)^2$ с помощью метода одномерной оптимизации. Полученное значение аргумента $y^* = 1.73$, подставляя его в уравнение связи, получим: $x^* = 2.615+0.923\cdot1.73 = 4.21$.

Далее, I_0 — число итераций алгоритма в случае применения модифицированного метода обратных вычислений для нахождения начальной точки; I_m , $I_{m \max}$ и $I_{m \min}$ — соответственно среднее, максимальное и минимальное число итераций в случае применения метода Монте-Карло; наконец, Р доля наблюдений, в которых число итераций оказывается больше при применении метода Монте— Карло, чем при применении модифицированного метода обратных вычислений. Можно обнаружить, что результат зависит от размеров интервалов: в случае, если границы интервалов располагаются близко к точке минимума, лучшие результаты наблюдаются при методе Монте-Карло, так как есть большая вероятность того, что сгенерированное случайное число окажется близким по значению с минимумом. Однако при увеличении интервалов с помощью модифицированного метода обратных вычислений можно найти точку минимума при гораздо меньшем числе итераций.

ЗАКЛЮЧЕНИЕ

На простых примерах формирования выручки приведено решение обратных задач модифицированным методом обратных вычислений. Основное преимущество данного метода заключается в простоте его компьютерной реализации. Рассмотрены различные варианты ограничений на значения показателей. Показана возможность применения модифицированного метода обратных вычислений для решения задач оптимизации, в частности, для определения начальной точки при оптимизации функции двух переменных методом Хука—Дживса. С увеличением интервалов модифицированный метод показал лучший результат, чем метод Монте-Карло. Рассмотренные методы могут быть по-

лезны в системах поддержки формирования управленческих решений.

Дальнейшие исследования будут направлены на решение многоаргументных задач заимствования, а также оптимизационных задач, в частности, предполагается разработка алгоритмов прямого поиска минимума многоаргументной функции с помощью модифицированного метода обратных вычислений.

ЛИТЕРАТУРА

- 1. Одинцов Б.Е., Романов А.Н. Проблемы создания информационных систем управления эффективностью бизнеса // Вестник Финансового университета. 2014. № 6. С. 22—36.
- 2. Медведев А.В., Победаш П.Н., Смольянинов А.В. Система поддержки принятия решений для управления региональным экономическим развитием на основе решения линейной задачи математического программирования // Международный журнал прикладных и фундаментальных исследований. 2013. № 12. С. 110—115.
- 3. Дик В.В. Методология формирования решений в экономических системах и инструментальные среды их поддержки. М.: Финансы и статистика, 2001. 300 с.
- Одинцов Б.Е. Обратные вычисления в формировании экономических решений. М.: Финансы и статистика, 2004. — 256 с.
- Виштак О.В., Штырова И.А. Использование технологии обратных вычислений при мониторинге качества дополнительного образования в ВУЗе // Вестник Астраханского государственного технического университета. 2014. № 2. С. 67—73.
- 6. *Бармина Е.А., Квятковская И.Ю.* Мониторинг качества коммерческой организации. Структурирование показателей. Применение когнитивных карт // Там же. 2010. № 2. С. 15—20
- 7. *Одинцов Б.Е.* Информационные системы управления эффективностью бизнеса. М.: Юрайт, 2015. 208 с.
- Грибанова Е.Б. Методы решения обратных задач экономического анализа // Корпоративные финансы. 2016. № 1. С. 119—130.
- 9. Одинцов Б.Е., Романов А.Н. Итерационный метод оптимизации управления предприятиями средствами обратных вычислений // Вестник Финансового университета. 2014. N 2. C. 60—73.
- Баканов М.И., Шеремет А.Д. Теория экономического анализа. М.: Финансы и статистика, 1997. 416 с.
- Моцкус Й.Б. О байесовских методах поиска экстремума // Автоматика и вычислительная техника. — 1972. — № 3. — С. 53—62.
- 12. *Мицель А.А., Шелестов А.А.* Методы оптимизации: учеб. пособие. Томск: ТУСУР, 2002. 17 с.

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

Грибанова Екатерина Борисовна — канд. техн. наук, доцент, Томский государственный университет систем управления и радиоэлектроники, ⊠ katag@yandex.ru.



УДК 336.76

СЕГМЕНТАЦИЯ И ХЭШИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ НА ФОНДОВОМ РЫНКЕ¹

А.Г. Спиро, М.Д. Гольдовская, Н.Е. Киселева, И.В. Покровская

Предложено каждому анализируемому временному ряду цены торгуемого на бирже актива (ВР-Ц) поставить в соответствие временной ряд хэш-кодов (ВР-ХК), которые для каждого элемента ВР-Ц будут показывать рост или падение цены. Отмечено, что в данном случае хэш-коды представляют собой целые числа, и их последовательность позволяет выделить в динамике изменения цены биржевого актива одинаковые (типовые) группы членов ряда ВР-Ц. Описаны процедуры преобразования исходного временного ряда и определения соответствующих хэш-кодов. Сформулированы основные свойства последовательности хэш-кодов. Предложена методика анализа траектории и прогнозирования цены биржевого актива по данным сегментации и хэширования.

Ключевые слова: фондовый рынок, процедура скользящего окна, хэш-коды, сегментация временных рядов, типовые сегменты, прогнозирование роста/падения котировок акций.

ВВЕДЕНИЕ

Профессиональные участники рынка используют в техническом анализе ряд различных индикаторов для прогнозирования цен [1, 2]. Поскольку принципы анализа достаточно подробно изложены в литературе, то обратим внимание лишь на тот факт, который сформулирован в одной из «аксиом» технического анализа, а именно, что в цене актива в данный момент времени отражены все возможные ситуации, которые известны участникам фондового рынка относительно данного актива. Это в свою очередь предполагает, что при одних и тех же внешних условиях (ситуациях) участники рынка будут вести себя одинаково. Заметим, что участники при работе на фондовом рынке располагают данными в виде ценовых графиков или временных рядов (ВР-Ц), отражающих цену торгуемого на бирже актива. При работе на фондовом рынке часто требуется анализировать не только входные данные, но и делать прогноз цены исследуемого актива (в статье рассматривается прогнозирование на ближайшее будущее). В настоящей работе для решения задач анализа и прогнозирования временных рядов на фондовом рынке предлагается метод, использующий процедуры хэширования [3, 4] и сегментации исследуемых временных рядов. Несколько слов о процедуре хэширования. Хэширование (от англ. hashing) — преобразование по заданному алгоритму входной последовательности чисел в выходную бинарную последовательность. Элементы бинарной последовательности, полученной из бинарной с помощью дополнительного преобразования, называют хэш-кодом или хэшем (подробнее см. далее § 1).

Основная идея этого метода состоит в следующем. Предлагается анализ исходной траектории заменить анализом последовательности (траектории) хэш-кодов, полученной из исходной траектории. Это сильно упрощает задачу анализа и прогнозирования, но благодаря тому, что теперь по последовательности хэш-кодов уже нельзя прогнозировать значение характеристики, а только ее знак, т. е. «увеличение» или «уменьшение» (а точнее, уменьшение или неизменность) значения этой характеристики, а именно такой прогноз и интересует большинство «игроков» фондового рынка. Поскольку хэш-коды в данном случае представля-

 $^{^1}$ Работа выполнена при частичной финансовой поддержке РНФ (грант 14-29-00309) и РФФИ (проекты 14-07-00463, 15-07-06713, 16-07-00895, 16-07-00896).



ют собой целые числа, то их последовательность позволяет выделить в динамике изменения цены биржевого актива одинаковые (типовые) группы членов (сегменты) ряда ВР-Ц. Анализируемые сегменты ряда ВР-Ц получаются с помощью скользящего окна заданного размера. Типовая группа это группа сегментов, для которой рост или падение котировочной цены в каждом сегменте происходит в одной и той же последовательности по времени. Другими словами, хэш-коды позволяют разделить исходный ряд ВР-Ц на сегменты одинаковой длины (равной размеру скользящего окна) и разделить все множество сегментов на типовые группы (в каждую типовую группу входят сегменты, имеющие одно и то же значение хэш-кода, которое и служит индексом типовой группы).

На этапе прогнозирования определяется типовая группа для последнего по времени сегмента временного ряда. Для каждого сегмента из этой типовой группы (всего таких сегментов V_0) определяется хэш-код для следующей по времени точки временного ряда относительно последней точки этого сегмента. Подсчитывается число сегментов V_2 , для которых такой хэш-код соответствует падению котировок. Тогда величина $P_{02} = V_2/V_0$ является оценкой значения вероятности падения котировок в следующий момент времени. По величинам P_{02} и V_0 можно рассчитать доверительный интервал значения такой вероятности, а значит, принять или отвергнуть гипотезу «в следующий момент времени будет падение котировок». Размер оценки значения вероятности роста котировок P_{01} определяется соотношением $P_{01} = V_1/V_0$, где V_1 число сегментов, для которых соответствующий хэш-код определяет рост котировок. На практике для принятия решения используется правило Байеса — по максимальному значению P_{0i} , i = 1, 2.

1. ПОСТАНОВКА ЗАДАЧИ

Необходимо осуществить сегментацию исходного ряда ВР-Ц для поиска типовых групп. Для этого используем прием, рассмотренный в работе [5].

- 1. На основе исходного временного ряда ВР-Ц строится двоичный кортеж (последовательность), состоящий из единиц и нулей. Значение «1» соответствует повышению или неизменности котировок в данный момент времени, а «0» понижению.
- 2. Для этой последовательности применяется пошаговая процедура скользящего окна шириной *m* с шагом одна позиция. Каждому окну, полученному в рамках этой процедуры, соответствует *m*-разрядный двоичный код.

- 3. Этот код преобразуется в десятичное целое число, которое и будет представлять собой хэш-код для соответствующего сегмента ряда ВР-Ц.
- 4. Полученная последовательность хэш-кодов и является искомым временным рядом хэш-кодов (BP-XK).
- 5. Используя BP-XK, выделяются и индексируются типовые группы сегментов ряда BP-Ц.

2. ПРЕОБРАЗОВАНИЕ ИСХОДНОГО ВРЕМЕННОГО РЯДА И ОПРЕДЕЛЕНИЕ ХЭШ-КОДОВ

Котировки биржевого актива являются временным рядом цены актива, элементы которого представляют собой динамическую переменную $C(t_j)$, которая наблюдается (измеряется) с некоторым постоянным шагом τ по времени, при этом выполняется условие: $t_j = t_0 + (j-1)\tau$, $C_j = C(t_j)$, $j = \overline{1,N}$, где N — объем выборки.

Определим BP-XK в результате преобразования ряда, который получен из исходного BP-Ц следующим образом:

$$X(t_{j}) = \begin{cases} 1, \text{ если } \Delta C(t_{j}) > 0 \\ 0, \text{ если } \Delta C(t_{j}) \le 0 \end{cases},$$

$$\Delta C(t_{j}) = C(t_{j}) - C(t_{j+1}). \tag{1}$$

Ряд (1) фактически представляет собой кортеж, компонентами которого служат 1 и 0. Для анализа кортежа (1) используется скользящее окно шириной $m \ll N$, где (N-1) — длина (размерность) кортежа (1). Такое окно позволяет ввести в рассмотрение двоичный позиционный код, образованный из единиц и нулей ряда (1). Старший разряд этого кода в момент времени t_j совпадает с двоичной переменной кортежа (1) $X(t_j)$, младший разряд совпадает, соответственно, с двоичной переменной $X(t_j - (m-1))$. Таким образом, в окне шириной m будет находиться двоичный позиционный код, имеющий взаимно однозначное соответствие с целым числом $d^m(T)$, для момента времени $t_i = T$

$$d^{m}(T) = \sum_{i=0}^{m-1} X(T-i)2^{m-i-1},$$
 (2)

где m — ширина скользящего окна; i = (0, 1, ..., m). Из формулы (2) непосредственно следует общий

вид
$$j$$
-го члена ряда (1): $d_j^m(T-j) = \sum_{i=0}^{m-1} X(T-i)$

 $-(j+i))2^{m-i-1}$, где j=0,1,...,N. Применяя процедуру «скользящего окна» для кортежа (1) и смещая окно на один шаг (одну позицию) от t=0 к

Таблица 1

Сегменты и их хэш-коды для ВР-Ц акций Сбербанка (за период с 20.02 по 05.03.2008 г.)

Дата		Сег	мент		ВР-Ц	Разность	Знак	Знач	чение хэш	-кода сегм	ента
дата	1	2	3	4	Б1-Ц	Тазноств	разности	разности 1		3	4
20.02					83,09	_	_	1	1	1	_
21.02	×				83,75	0,66	1		_	_	_
22.02	×	×			82,99	-0,76	0			_	_
26.02	×	×	×		85,10	2,11	1	5		1	_
27.02	×	×	×	×	83,78	-1,32	0	3	2		
28.02	×	×	×	×	82,95	-0,83	0		2		
29.02	×	×	×	×	80,20	-2,75	0			1	32
03.03		×	×	×	78,70	-1,50	0	_			32
04.03			×	×	76,50	-2,20	0	_	_		
05.03				×	78,32	1,82	1	_	_	_	

t=T, получим последовательность целых чисел, которые и являются хэш-кодами окон, соответствующих сегментам анализируемого BP-Ц. В качестве примера в табл. 1 приведены значения BP-Ц для ПАО «Сбербанк», разность цен $\Delta C(t_j)$, вычисленная для смежных дней на основе BP-Ц, знак этой разности, кортеж из единиц и нулей, характеризующий знак разности, и хэш-коды, которые рассчитаны для четырех шестидневных сегментов. Будем считать, что местоположение старшего разряда двоичного позиционного кода, который определяет хэш-код, всегда соответствует последней дате шестидневного сегмента. Заметим, что для скользящего окна шириной m=6 будет 64 хэш-кода, т. е. 64 различных типовых сегмента.

3. ИСПОЛЬЗОВАНИЕ ХЭШ-КОДОВ ДЛЯ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ

Исходные ВР-Ц и ВР-ХК позволяют совместно провести анализ траектории движения цены биржевого актива. Предлагается методика такого анализа, суть которой состоит в следующем.

- 1. В распоряжении инвестора всегда имеются данные о котировочной цене биржевого актива за интересующий интервал времени вплоть до текущего момента T.
- 2. Динамику движения цены отражает существующий ВР-Ц на всех интервалах (T-j), где j=0,1,...,N-1,N- объем выборки.
- 3. На всех интервалах, кроме начального, вместе с ценой известно значение хэш-кода (элемент ВР-ХК). Другими словами, между сегментами

ВР-Ц и ВР-ХК существует соответствие — каждому сегменту ВР-Ц соответствует элемент ВР-ХК.

4. В текущий момент времени T всегда известен хэш-код $d_0^m(T)$, который определим как *текущий хэш-код*, а сегмент ВР-Ц, соответствующий текущему хэш-коду, определим как *текущий сегмент* ВР-Ц. Текущий хэш-код является последним по времени хэш-кодом в выборке.

Методика состоит из двух этапов — набора статистики и прогнозирования.

На первом этапе набираются биржевые данные, которые используются для получения статистики частоты вхождения в выборку типовых сегментов и хэш-кодов. С этой целью для каждого типа хэш-кода определяются два параметра. Первый — число вхождений в выборку самого хэш-кода и второй — это число вхождений в выборку пары, состоящей из хэш-кода данного типа и хэш-кода, непосредственно следующего за ним, например, при падении котировок.

В начале второго этапа (прогнозирование) последний в выборке хэш-код рассматриваемого типа принимается искусственно за текущий и относительно него осуществляется прогноз. Правильность прогноза, т. е. значение прогнозируемого хэш-кода, проверяется на имеющейся выборке. Таким образом, вначале все хэш-коды разделяются на два класса: первый, у которого результаты прогноза совпали с историей котировок, и второй, — там, где этого не произошло. По численности первого и второго класса можно сделать вывод о целесообразности применения этого метода прогнозирования.



Затем начинает работать процедура прогнозирования для текущего хэш-кода. Далее приведено более детальное ее описание.

- 1. Прогнозирование движения цены (рост или падение) сводится к выделению в последовательности BP-XK всех хэш-кодов, равных текущему $d_0^m(T)$.
- 2. После этого необходимо выделить пары хэшкодов $d_0^m(T)$, $d_2^m(T+1)$, следующие за $d_0^m(T)$ (индекс 2 указывает на падение или неизменность котировок, а индекс 1 — соответствует их росту).
- 3. Определить наиболее вероятный хэш-код из этой пары можно с помощью следующего алгоритма.
- 3.1. Определим число вхождений V_0 хэш-кода $d_0^m(T)$ в ряд BP-XK.
- 3.2. Аналогично п. 3.1 определим число вхождений V_2 для пары хэш-кодов, непосредственно следующих один за другим $d_0^m(T)$, $d_2^m(T+1)$.
- 3. Оценим вероятность P_{02} наступления события $d_2^m(T+1)$:

$$P_{02} = V_2 / V_0. (3)$$

4. Оценим вероятность наступления события $d_1^m(T+1)$:

$$P_{01} = 1 - P_{02}. (4)$$

Соотношения (3) и (4) позволяют на основе имеющихся статистических данных сделать вывод о возможном направлении движения цены актива: если $P_{01} > P_{02}$, то выбираем хэш-код «рост цены»; если $P_{01} \le P_{02}$, то выбираем хэш-код «падение цены».

4. ПРОВЕРКА МЕТОДИКИ НА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Разработанная методика проверялась на данных как для срочного рынка FORTS биржи РТС, так и для рынка акций ММВБ.

В табл. 2 представлены фрагменты ВР-Ц (индекс РТС) и ВР-ХК, полученного для скользящего окна шириной m=6.

В табл. 3 представлен результат обработки массива данных индекса РТС с шагом по времени (таймфреймом), равном одному дню (период с 29.11.2011 по 29.01.2013 г.) для хэш-кода 25. Как было сказано ранее, каждому хэш-коду можно поставить в соответствие типовую группу сегментов, полученных с помощью процедуры скользящего окна. Тип этого сегмента определяется хэш-кодом

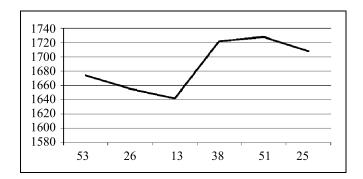


Рис. 1. Типовой сегмент 1, 28.02.2012 г., хэш-код 25 индекса РТС

последнего по времени элемента ВР-Ц, входящего в окно. На рис. 1 показан типовой сегмент хэшкода 25 для окна шириной m=6 (в табл. 3 это третья строка).

По оси ординат указано значение индекса РТС, по оси абсцисс — последовательность хэш-кодов. Последним по времени является хэш-код 25. Этот хэш-код и определяет типовую группу. Пример другого сегмента из этой же группы показан на рис. 2 (в табл. 3 — это вторая строка).

Визуально эти сегменты отличаются друг от друга, однако общим признаком для этой типовой группы (как и для всех остальных типовых групп) служит рост или падение котировочной цены при

 $\it Taблицa~2$ Фрагменты ВР-Ц и ВР-ХК

Дата	Индекс РТС	Хэш-код		
17.10.2012	1513,96	50		
18.10.2012	1512,01	25		
19.10.2012	1494,44	12		
22.10.2012	1497,63	38		
23.10.2012	1456,73	19		
24.10.2012	1462,43	41		
25.10.2012	1456,91	20		
26.10.2012	1441,38	10		
29.10.2012	1435,05	05		

Таблица 3 Выборка с использованием хэш-кода **25**

Дата	Индекс РТС	Хэш-код		
21.01.2013	1600,13	25		
18.10.2012	1512,01	25		
28.02.2012	1708,16	25		
29.11.2011	1466,36	25		



переходе от одной временной точки в пределах окна (сегмента) к другой. Так, например, для окна ширины 6 всего будет 64 различные типовые группы, которые можно пронумеровать от 0 до 63, и в каждой такой группе одноименные позиции будут «вести» себя одинаково (цены будут расти или падать), но при этом сами размеры роста или падения будут различаться. Максимальное число сегментов $R_{\rm max}$ при ширине окна m и выборке N равно N-m-1.

Данная методика сегментации и прогнозирования была проверена на акциях Газпрома. Анализируемый период составлял 1416 дней с 10.12.2007

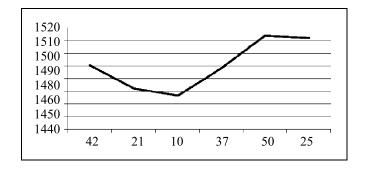


Рис. 2. Типовой сегмент 2, 18.10.2012 г., хэш-код 25 индекса РТС

Tаблица 4 Результаты прогнозирования движения котировок акций Газпрома за период с 10.12.2007 по 16.08. 2013 г.

1	2	3	4	5	6	7	1	2	3	4	5	6	7
0	17	6	0,35	0,65	1		33	21	8	0,38	0,62	1	
1	22	10	0,45	0,55	1		34	25	16	0,64	0,36	1	
2	19	9	0,47	0,53	1		35	29	18	0,62	0,38	1	
3	23	12	0,52	0,48	1		36	23	13	0,57	0,43	1	
4	26	11	0,42	0,58		1	37	23	12	0,52	0,48		1
5	18	7	0,39	0,61	1		38	27	14	0,52	0,48	1	
6	30	15	0,50	0,50			39	17	6	0,35	0,65	1	
7	22	7	0,32	0,68		1	40	24	6	0,25	0,75	1	
8	26	16	0,62	0,38	1		41	24	11	0,46	0,54	1	
9	23	9	0,39	0,61		1	42	14	5	0,36	0,64	1	
10	24	8	0,33	0,67	1		43	18	14	0,78	0,22	1	
11	18	10	0,56	0,44	1		44	20	7	0,35	0,65		1
12	32	15	0,47	0,53	1		45	28	14	0,50	0,50		
13	26	15	0,58	0,42	1		46	16	10	0,63	0,38	1	
14	16	10	0,63	0,38	1		47	11	2	0,18	0,82		1
15	24	11	0,46	0,54		1	48	27	14	0,52	0,48		1
16	15	7	0,47	0,53		1	49	19	8	0,42	0,58	1	
17	35	18	0,51	0,49		1	50	19	10	0,53	0,47		1
18	26	12	0,46	0,54		1	51	23	16	0,70	0,30	1	
19	21	10	0,48	0,52	1		52	30	19	0,63	0,37	1	
20	18	10	0,56	0,44	1		53	12	6	0,50	0,50		
21	20	13	0,65	0,35	1		54	27	15	0,56	0,44	1	
22	22	9	0,41	0,59		1	55	14	11	0,79	0,21		1
23	13	7	0,54	0,46		1	56	23	11	0,48	0,52	1	
24	23	14	0,61	0,39		1	57	15	7	0,47	0,53	1	
25	28	16	0,57	0,43		1	58	16	6	0,38	0,63	1	
26	26	13	0,50	0,50			59	13	6	0,46	0,54	1	
27	28	12	0,43	0,57		1	60	19	7	0,37	0,63	1	
28	19	9	0,47	0,53	1		61	18	7	0,39	0,61	1	
29	13	6	0,46	0,54		1	62	20	12	0,60	0,40	1	
30	15	8	0,53	0,47		1	63	11	7	0,64	0,36	1	
31	20	15	0,75	0,25	1						Итого	40	20
32	22	7	0,32	0,68	1						711010	70	20



по 16.08.2013 г. В качестве основного таймфрейма был принят день. Объем выборки составлял 1416 значений. В табл. 4 приведены результаты прогнозирования.

Результаты анализа и прогнозирования роста и падения котировок акций с использованием хэшкодов, представленных в табл. 4, можно интерпретировать следующим образом.

- Предложенный алгоритм прогнозирования носит вероятностный характер, и на примере акций Газпрома за указанный период вероятность совпадения реального роста и падения котировок с прогнозом составляет 0,66, что по оценкам экспертов-брокеров является достаточно высоким результатом, поскольку при прогнозировании не использовалось никакой другой информации, кроме временного ряда цен исследуемого актива (инсайдерской информации, цен на нефть, золото или другие базовые активы, биржевые индексы и др.).
- Вопреки устоявшемуся мнению, что значения вероятностей роста и падения должны быть близки к 0,5, из табл. 4 следует, что более 42 % значений равны или превышают 0,6 (в табл. 4 они помечены полужирным шрифтом). Результат 0,5 (значения вероятностей роста и падения равны) получены только для хэш-кодов 6, 26, 45 и 53.
- По данным, для которых значения соответствующих вероятностей равны или превышают 0,6, совпадение реального движения цены с прогнозом составляет около 80 %.

По столбцам в табл. 4 указано: 1 — все хэш-коды, полученные в последовательности BP-XK; 2 — число вхождений V_0 каждого хэш-кода в BP-XK за указанный период; 3 — число вхождений V_1 пары из двух хэш-кодов: хэш-код, указанный в первом столбце и хэш-код смежный с ним при падении цены; 4 — значение оценки вероятности при падении цены (см. формулу (3)); 5 — значение оценки вероятности при росте цены (см. формулу (4)); 6 — результат прогноза и реальное движение цены за указанный период совпали; 7 — результаты не совпали.

ЗАКЛЮЧЕНИЕ

Предложен новый метод прогнозирования на фондовом рынке, базирующийся на процедурах сегментации и хэширования соответствующих временных рядов. Благодаря использованию хэш-кодов удается существенно упростить задачу прогноза — прогнозируется не значение характеристики, а только ее знак, т. е. «увеличение» или «умень-

шение» значения этой характеристики, а именно такой прогноз и интересует большинство «игроков» фондового рынка. Поскольку сами хэш-коды представляют собой целые числа, то их последовательность позволяет выделить в динамике изменения цены биржевого актива одинаковые (типовые) группы членов временного ряда (сегменты). Анализируемые сегменты временного ряда цены торгуемого на бирже актива получаются с помощью процедуры скользящего окна заданного размера. Типовая группа — это группа сегментов, для которой рост или падение котировочной цены в каждом сегменте происходит в одной и той же последовательности во времени. Другими словами, хэш-коды позволяют разделить исходный временной ряд цены актива на сегменты одинаковой длины (равной размеру скользящего окна) и разделить все множество сегментов на типовые группы.

Прогнозирование осуществляется на базе информации о хэш-кодах сегментов типовой группы для последнего по времени сегмента временного ряда.

Разработанная методика проверялась на данных как для срочного рынка FORTS биржи РТС (индекс РТС), так и для рынка акций ММВБ (акции Газпрома). Полученные результаты подтвердили эффективность предложенного метода.

ЛИТЕРАТУРА

- 1. *О'Нил У.Джс.* Как делать деньги на фондовом рынке. Стратегия торговли на росте и падении. М.: Изд. дом «Альпина», 2003. 328 с.
- Толли Т. Игра на понижение или техника «коротких» продаж. Правила игры финансовых топ-менеджеров Уоллстрит на фондовом рынке. М.: Торговый дом «Грант», 2004. 366 с.
- Ахо А., Хопкрофт Д. Структуры данных и алгоритмы. М.: Изд. дом «Вильямс», 2000. — С. 116—127.
- Кормен Т.Х., Лейзерсон Ч.И., Ривест Р.Л., Штайн К. Алгоритмы: Построение и анализ: 2-е изд. М.: Изд. дом «Вильямс», 2005. С. 282—315.
- Спиро А.Г., Дорофеюк Ю.А. Структурно-графовой подход к анализу фондового рынка // Проблемы управления. — 2011. — № 6. — С. 61—65.

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

Спиро Арнольд Григорьевич — канд. техн. наук, ст. науч. сотрудник, \bowtie arn.spi@mail.ru,

Гольдовская Марина Дмитриевна — науч. сотрудник, \bowtie mdgold54@mail.ru,

Киселева Нелли Евсеевна — науч. сотрудник, ⊠ lab55@ipu.ru,

Покровская Ирина Вячеславовна — науч. сотрудник, ⊠ ivp750@mail.ru,

2 11p7300 man.ra,

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

УДК 004.75:004.056.3

ОПТИМАЛЬНОЕ ОПЕРАТИВНОЕ РЕЗЕРВИРОВАНИЕ ИНФОРМАЦИИ В СИСТЕМАХ ОБРАБОТКИ ДАННЫХ НА БАЗЕ ВЫЧИСЛИТЕЛЬНЫХ СЕТЕЙ

Е.А. Микрин, С.К. Сомов

Рассмотрены вопросы применения оперативного резервирования информации в распределенных системах обработки данных (РСОД), функционирующих в рамках вычислительных сетей, в целях повышения безопасности их работы. Проанализированы особенности применения трех основных стратегий оперативного резервирования в РСОД. Сформулированы задачи оптимального по разным критериям распределения оперативного резерва РСОД по узлам вычислительной сети и предложены методы их решения. Рассмотрены проблемы и задачи оптимизации резервирования массивов данных с учетом их взаимосвязи между собой и резервирование данных в распределенных базах данных. Выполнен анализ влияния трафика запросов на распределение резерва РСОД.

Ключевые слова: распределенные системы обработки данных, вычислительные сети, резервирование массивов данных, задачи оптимизации резервирования, стратегии резервирования.

ВВЕДЕНИЕ

Распределенные системы обработки данных (РСОД) представляют собой сложные аппаратно-программные комплексы или системы, компоненты которых распределены на больших расстояниях друг от друга и объединены в единую систему различными каналами связи и специальным системным программным обеспечением. Они строятся, как правило, на основе глобальных, корпоративных или сетей специального назначения.

В процессе работы РСОД могут происходить инциденты, приводящие к появлению ошибочных данных, ошибочным результатам решения задач и обработки запросов системой, а в крайнем случае и к невозможности нормального функционирования РСОД, т. е. к потере системой работоспособности, ее отказу. Разрушение программного и информационного обеспечения системы может потребовать значительных затрат времени и ресурсов для восстановления потерянной в результате инцидента информации и восстановления работоспособности системы. А в некоторых случаях восстановление информации и, соответственно, работоспособности системы становится просто невозможным.

Примером воздействия ошибки в программном обеспечении на работу и безопасность сложной системы может служить катастрофа, произошедшая с первым космическим аппаратом «Mariner-1». Данный аппарат в соответствии с программой «Маринер» должен был стартовать 22 июля 1962 г. и лететь в сторону Венеры [1]. Однако примерно через 5 мин после старта произошла авария, и аппарат был разрушен. В свидетельских показаниях, полученных во время послеполетного расследования инцидента, указывалось, в частности, что в результате ошибки в программном обеспечении управляющие подсистемы аппарата получали неверно интерпретированную информацию. В итоге компьютер аппарата дал серию ошибочных команд на корректировку курса, в итоге аппарат потерял управление и потерпел крушение [2, 3].

Иллюстрацией того, к каким непроизводительным затратам времени приводит разрушение информации может быть инцидент, произошедший в OCLC Research Library Partnership (некоммерческий компьютерный библиотечный сервис и научно-исследовательская организация, цель которых заключается в предоставлении широкого доступа к различной информации в режиме он-лайн) [4]. Эта организация предоставляет доступ к информации с помощью RLIN (компьютерная информаци-



онная сеть библиотек). В этой информационной сети произошла авария, в результате которой был потерян большой объем данных. На восстановление разрушенной информации было потрачено 12 дней кропотливой работы ИТ-специалистов организации [5].

Причины, вызывающие возникновение инцидентов в процессе работы РСОД, приводящих к потере, разрушению или искажению информации и, в итоге, к нарушению безопасности и даже к отказу в работе системы, можно классифицировать следующим образом.

- Сбои и ошибки в работе оборудования вычислительной системы (BC), на базе которой работает РСОД: сбои в каналах связи, перебои в системах электропитания, сбои в системах хранения информации, сбои в работе вычислительных машин (серверов, рабочих станций).
- Некорректная работа обслуживающего персонала и пользователей системы.
- Потери информации в системах архивации в результате сбоев оборудования.
- Ошибочные данные и потеря информации при некорректной работе программного обеспечения, содержащего ошибки в коде.
- Заражение системы компьютерными вирусами.
- Потери и искажение данных, вызванные несанкционированным доступом в систему.

Для повышения безопасной работы распределенных систем обработки данных вводят различного рода избыточность [5]. В таких системах задача обеспечения сохранности используемых программ и данных решается главным образом применением информационной избыточности (созданием и хранением резервных данных). Кроме того, с этой же целью применяется и аппаратная избыточность (использование резервных устройств хранения данных), а также временная избыточность — использование части ресурсов системы для выполнения процессов дампирования (создания копий данных) и восстановления данных в случае их разрушения. Основным методом обеспечения сохранности данных в РСОД при ее функционировании служит метод резервирования, который предполагает использование следующих видов избыточности: идентичной избыточности (создания и хранения копий данных) и/или неидентичной (создания и хранения предысторий данных).

1. ПОСТАНОВКА ЗАДАЧИ

В настоящее время в системах обработки данных (СОД) различного класса и назначения применяются два вида резервирования: оперативное и восстановительное.

При оперативном резервировании некоторая часть информационных ресурсов (резервные ко-

пии и/или предыстории массивов данных) РСОД и ее вычислительной мощности используется только для поддержания работоспособности системы в случае, если при решении РСОД текущих задач произошло разрушение используемого ей массива/массивов данных.

В отличие от оперативного резервирования в восстановительном резервировании часть ресурсов системы используется только для восстановления используемых системой рабочих/текущих копий и/или предысторий массивов данных в случае их разрушения.

При оперативном резервировании данных в СОД используются три основные стратегии оперативного резервирования [5, 6]. Первая стратегия применяется для массивов постоянных данных (к этой категории можно отнести относительно редко меняющиеся программные модули или, например, справочники, содержащие статичную информацию разного рода). Вторая и третья стратегии применяются к текущим массивам оперативных данных, содержащим данные, которые подвергаются частым изменениям (добавление и удаление данных, изменение данных, их сортировка), например, таблица с проводками по аналитическим счетам в автоматизированной учетной системе банка или программный модуль на этапе его тестирования и исправления обнаруженных ошибок.

Стратегия I заключается в использовании некоторого числа копий основного массива данных. В случае, если основной массив разрушен, то для продолжения решения задачи СОД переходит к использованию первой его копии, если же и она разрушена, то используется следующая копия и т. д.

Стратегия II основывается на учете особенностей обновления текущих массивов данных и основывается на том, что вместо копий текущего массива используются его предыстории (предыдущие версии массива и журналы их изменений). В случае разрушения текущего массива запускается специальная программа, которая восстанавливает разрушенный массив на основе его предыдущей версии с использованием записей журнала изменений. Если восстановление закончилось неудачно, то восстановление текущего массива производится из следующей предыстории и т. д.

Стратегия III — это смешанная стратегия, которая использует возможности первых двух стратегий, т.е. эта стратегия использует и копии и предыстории текущего массива данных. Причем, сначала в соответствии со стратегией I используются копии массива, а затем, в случае неудачи восстановления из копий, используются предыстории массива в соответствии со стратегией II.

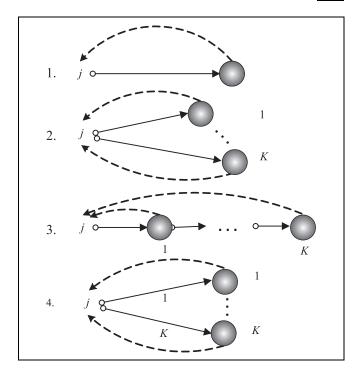
Объединение автономных компьютеров посредством каналов связи позволило получить единую распределенную BC, обладающую качественно



новыми возможностями по созданию и эксплуатации распределенных СОД по сравнению с возможностями систем, работающих на базе отдельных компьютеров. Применение методов резервирования в таких ВС по сравнению с автономными ЭВМ обладает рядом особенностей [7]:

- при анализе эффективности и выборе оптимальной стратегии резервирования данных в РСОД, построенной на базе ВС, необходимо учитывать параметры самой ВС: топологию сети, задержку сообщений в каналах связи, характеристики производительности отдельных компьютеров, надежность и пропускную способность объединяющих их каналов связи и т. д.;
- в силу территориальной распределенности в вычислительной сети можно использовать один из следующих основных способов хранения резерва РСОД:
 - централизованное, при котором резерв хранится в одном из узлов ВС;
 - децентрализованное, когда резерв распределен по некоторому множеству узлов ВС;
- в случае децентрализованного хранения резерва в РСОД для обработки поступающих в систему запросов можно применять несколько дисциплин обработки. Рассмотрим основные из них (см. рисунок) [7, 8].
- 1. Запрос пересылается для обработки в ближайший (согласно некоторому критерию) узел, который хранит резерв массива данных, необходимого для обработки запроса.
- 2. Запрос пересылается для обработки параллельно в *K* узлов с требуемым резервом.
- 3. Запрос, который изначально был получен в j-м узле, последовательно пересылается по узлам пути длины K, начинающегося в j-м узле и состоящим из узлов, содержащих резерв требуемого массива. Запрос последовательно пересылается по узлам данного пути до тех пор, пока он или не будет успешно обработан в некотором узле либо не будет пройден весь путь из K узлов.
- 4. Запрос, полученный в j-м узле, поочередно отправляется для обработки в K ближайших узлов, содержащих необходимый резерв. Опрос узлов продолжается до тех пор, пока из очередного опрашиваемого узла не будет успешно получен ответ на запрос либо не будут опрошены все K узлов без получения требуемого ответа.

Территориальная распределенность ВС обусловливает не только возможность применения различных дисциплин обработки запросов, но и большое разнообразие вариантов размещения резерва по узлам сети, которые обеспечивают различные значения стоимости, времени и надежности обработки поступающих запросов [7]. Эти обстоятельства определяют необходимость решения задачи оптимального размещения оперативного



Основные дисциплины обработки запросов

резерва массивов данных по узлам ВС в соответствии с различными критериями оптимизации. В результате решения этой задачи определяется множество $I^{\rm o}=\{i|y_i=1\}$ узлов сети, в которых необходимо разместить резерв $(y_i=1,$ если в i-м узле размещен резерв), а также количество копий и/или предысторий массива данных, размещаемых в узлах этого множества $I^{\rm o}$. В качестве критериев оптимизации решения данной задачи можно, например, использовать такие критерии, как максимум вероятности и минимум среднего времени получения ответа на запрос, минимум стоимости обработки запросов и минимум затрат на хранение резерва.

Например, задача поиска размещения оперативного резерва в узлах BC, обеспечивающего минимум стоимостных затрат S(X) на обработку поступающих запросов и хранение резерва формулируется таким образом:

$$S(X) \to \min$$
 (1)

при ограничениях на время t_j обработки запросов, вероятность ρ_j получения ответа на запрос и на размер x_j резерва в узлах сети:

$$t_j(X) \le \overline{t}, \quad j \in J_3,$$
 (2)

$$\rho_j(X) \ge \bar{\rho}_j, \quad j \in J_3,$$
(3)

$$x_j \le \bar{x}_j, \quad j \in I^{\circ}.$$
 (4)



Здесь J_3 — это множество индексов узлов BC, в которых возникают запросы; $I^\circ = \{i|y_i=1\}$ — множество узлов BC, в которых размещен резерв; $X=(x_j)$ — распределение резерва по узлам сети, где x_j — число копий и/или предысторий, размещенных в узле j; \bar{x}_j — максимальное число копий и/или предысторий массива данных, которые можно разместить в узле j.

Размер затрат S(X) для дисциплин 1 и 2 обработки запросов определяется по формуле

$$S(X) = L \sum_{i \in I^{\circ}} s_i x_i + \sum_{j=1}^{N} \sum_{i \in I^{\circ}} ZP_{ji}(x_i) (U_j \psi_{ji} + V_j).$$

Здесь $\psi_{ji}\in \Psi^{\mathrm{o}};$ матрица $\Psi^{\mathrm{o}}=\|\psi_{ji}^{\mathrm{o}}\|,$ где $\psi_{ji}^{\mathrm{o}}\in\{0,\,1\}$

и $\psi_{ji}^{o} = 1$, если запрос из j-го узла адресуется в i-й узел, s_i — стоимость хранения одного бита информации в узле i; L — длина массива данных в битах; $U_j(V_j)$ — интенсивность информационных запросов (запросов на модификацию), возникающих в узле j; $ZP_{ji}(x_i)$ — размер средних производительных затрат на обработку запроса, посланного из узла j на обработку в узел i:

$$ZP_{ii}(x_i) = 2D_{ii} + E_i(x_i)h_i,$$

где $E_i(x_i)$ — среднее время обработки запроса в узле i, D_{ji} — стоимость передачи запроса/ответа из узла j в узел i (предполагается, что $D_{ji} = D_{ij}$), h_i — стоимость использования ресурсов ЭВМ узла i в единицу времени.

Значения времени обработки запросов и вероятности получения ответа на запрос определяются по формулам

для дисциплины 1:

$$t_{j}(X) = 2t_{3}(1 - \psi_{jj}) + \sum_{k=1}^{N} E_{k}(x_{k})\Psi_{jk};$$
$$\rho_{j}(X) = \sum_{k=1}^{N} r_{jk} P_{k}(x_{k})\Psi_{jk},$$

для дисциплины 2:

$$t_{j}(X) = \max \left\{ 2t_{3} + \max E_{k}(x_{k}); E_{j}(x_{j})\Psi_{jj} \\ k \in N_{j} \\ k \neq j \right\},$$

$$\rho_{j}(X) = 1 - \prod_{k=1}^{N} [1 - r_{jk} P_{k}(x_{k}) r_{kj} \Psi_{jk}],$$

где $E_k(x_k)$ — среднее время обработки запроса в узле $k, P_k(x_k)$ — вероятность успешной обработки запроса в узле k при условии размещения в этом узле резерва объемом x_k .

2. РЕШЕНИЕ ЗАДАЧИ

В работе [7] показано, что сформулированная в настоящей статье задача поиска оптимального оперативного резерва принадлежит классу задач целочисленного выпуклого программирования. Для этого в указанной работе доказывается выпуклость ограничений (2)—(4) и целевой функции (1).

Достаточное условие совпадения локального и глобального минимумов целевой функции заключается в выпуклости функций S(X), $t_j(X)$ и $\rho_j(X)$. Данное обстоятельство позволяет применить для решения нашей задачи оптимизации оперативного резервирования методы, применяемые для решения задач распределения ограниченных ресурсов [9].

Аналогично можно показать, что к классу задач целочисленного выпуклого программирования принадлежат и задачи оптимального оперативного резервирования, приведенные ниже и использующие в качестве критериев оптимизации:

- максимум вероятности успешной обработки запросов (задача (5)—(9));
- минимум среднего времени получения ответа на запрос (задача (9)—(12)).

Формулировки данных задач оптимального оперативного резервирования:

Найти:

$$\max R(X) = \max \prod_{j \in J_2} \rho_j(X) \tag{5}$$

при ограничениях:

$$S(X) \le \bar{S} \,, \tag{6}$$

$$t_i(X) \le \bar{t} \ , \quad j \in J_3, \tag{7}$$

$$x_i \le \bar{x}_i, \quad j \in I^{\circ}.$$
 (8)

Найти:

$$\min T(X) = \min N_3^{-1} \sum_{i \in J} t_i(X),$$
 (9)

где $N_3 = |J_3|$,

при ограничениях:

$$S(X) \le \overline{S} \,, \tag{10}$$

$$\rho_{j}(X) \ge \overline{\rho} \,, \quad j \in J_{3}, \tag{11}$$

$$x_i \le \bar{x}_i, \quad j \in I^{\circ}.$$
 (12)

Приведенные задачи оптимального оперативного резервирования массивов данных в ВС сформулированы в предположении, что они решаются в два этапа. На первом этапе определяется оптимальное размещение копий массива по узлам ВС $I^{\rm o}$ (не более одной копии в каждом узле). Это раз-



мещение получается в результате решения задачи (1)—(4) с помощью известных алгоритмов [10—14].

Если найденное на первом этапе размещение копий не обеспечивает выполнение ограничения (3) задачи, т. е. не обеспечивает требуемую вероятность обработки запросов $\bar{\rho}$, то переходят ко второму этапу решения задачи. На втором этапе для каждого узла множества $I^{\rm o}$ решается задача определения оптимального объема резерва в отдельном узле ВС. Для ее решения применяются традиционные методы (методы решения задач распределения ограниченных ресурсов) [7].

Решение задачи оптимального оперативного резервирования массивов данных в ВС в два этапа позволяет уменьшить вычислительную сложность задачи, так как вместо одной сложной задачи решаются две более простые задачи. Решение первой из них позволяет определить подмножество узлов ВС, в которых необходимо разместить резерв. А решение второй задачи позволяет определить объем резерва, размещаемого в полученном ранее подмножестве узлов ВС.

Рассмотрим более общую постановку задачи оптимального оперативного резервирования, в которой могут использоваться аналогичные критерии оптимизации.

Рассмотрим такую постановку задачи, например, для дисциплины 1 обработки запросов при использовании в качестве критерия оптимизации минимума стоимости обработки запросов и стоимости хранения резерва. Задача формулируется следующим образом:

найти минимум значения функции

$$S(X) = L \sum_{i \in I^{0}} s_{i} x_{i} + \sum_{j=1}^{N} V_{j} \sum_{i/x_{i} \neq 0} [2D_{ji} + E_{i}(x_{i})h_{i}] + \sum_{j=1}^{N} U_{i} \Psi_{ji} [2D_{ji} + E_{i}(x_{i})h_{i}]$$

при ограничениях:

$$\begin{split} \rho_j(X) &= \rho_{jk} \geq \overline{\rho} \;, \quad j \in J_3; \quad t_j(X) = t_{jk} \leq \overline{t} \;, \quad j \in J_3; \\ x_i &\in \{0, \, 1, \, ..., \, \overline{x}_i\}, \quad j = \overline{1, \, N} \;, \end{split}$$

где k — индекс узла сети с оперативным резервом, для которого справедливы условия:

$$2D_{jk}+E_k(x_k)h_k=\min_{i\;\in\;N_j}[2D_{ji}+E_i(x_i)];$$
 $N_j=\{i/
ho_{ji}\leqar
ho\;,\;t_{ji}\leqar t\;\};$ $ho_{ji}=r_{ji}P_i(x_i)r_{ij}$ при $i\neq j$ и $ho_{ji}=P_i(x_i)$ при $i=j;$ $t_{ji}=2t_3+E_i(x_i)$ при $i\neq j$ и $t_{ji}=E_i(x_i)$ при $i=j,$

где $E_k(x_k)$ — среднее время обработки одного запроса в узле $k;\ t_3$ — среднее время задержки пере-

дачи сообщений в каналах связи сети; h_k — стоимость использования ресурсов компьютера узла k в единицу времени; D_{jk} — стоимость передачи сообщения (запроса/ответа на запрос) из узла j в узел k (полагаем, что $D_{jk} = D_{kj}$).

Аналогично формулируются и другие задачи оптимального оперативного резервирования, использующие в качестве критерия оптимизации минимум среднего времени получения ответа на запрос или максимум вероятности успешной обработки запросов.

3. РЕЗЕРВИРОВАНИЕ ВЗАИМОСВЯЗАННЫХ МАССИВОВ ДАННЫХ

Задачи оптимального оперативного резервирования (см. § 1 и 2) были сформулированы в предположении, что резервируемые в РСОД массивы данных используются независимо друг от друга. Это позволило ставить и решать задачи оптимального резервирования отдельно и независимо для каждого массива данных. Однако чаще всего на практике в СОД разного класса и назначения, в том числе и в РСОД, построенных на базе ВС, решаются задачи и обрабатываются запросы, работа с которыми требует использования некоторого подмножества массивов данных. Причем, как правило, в данное подмножество входят массивы данных, взаимосвязанных между собой (ссылками или алгоритмически). В частности, обработка одного запроса к распределенной базе данных (РБД) также может потребовать доступ к данным из нескольких связанных друг с другом массивов данных, которые могут быть размещены в разных узлах ВС, территориально удаленных друг от друга.

По этой причине возникает необходимость оптимизации резервирования массивов данных, используемых в РСОД, с учетом их взаимосвязи друг с другом и других параметров ВС (таких, как топология сети, трафик запросов различного типа, надежность функционирования каналов связи и отдельных ЭВМ сети и др.). Покажем, что данные задачи можно сформулировать, используя приведенные выше результаты.

Рассмотрим работу РСОД, построенную на основе ВС, состоящей из N узлов. В системе используется M массивов данных и решается J различных задач. Каждая из задач системы может использовать несколько различных массивов и может решаться в нескольких различных узлах сети. Обозначим через W_j то число решений j-й задачи, $j=\overline{1,J}$, которое производится за некий интервал времени (например, час, сутки, неделя, месяц, квартал и т. п.).



Введем переменную $\Psi_{jn} \in (0, 1)$ такую, что $\Psi_{jn} = 1$, если j-я задача решается в n-м узле сети; обозначим U_{jm} — среднее число информационных запросов j-й задачи к m-му массиву данных, а V_{jm} — среднее число запросов на модификацию m-го массива данных, возникающих при решении j-й задачи; x_{nm} — объем резерва m-го массива, размещенного в n-м узле сети, L_m — размер m-го массива данных в байтах.

Условимся, что задача в системе решена успешно только в случае, если получены ответы на все запросы, возникающие в процессе ее решения (запросы на получение и модификацию данных) ко

всем требуемым массивам данных. Тогда вероятность $P_{nj}(X)$ решения j-й задачи в n-м узле при распределении резерва $X = \|x_{nm}\|$

$$P_{nj}(X) = \prod_{m=1}^{M} [\rho_{nm}(X)]^{U_{jm}} \beta_{nm}^{V_{jm}},$$

где $\rho_{nm}(X)$ — вероятность получения ответа на информационный запрос к m-му массиву, выданный в n-м узле сети; $\beta_{nm}(X)$ — это вероятность успешной обработки выданного в n-м узле сети запроса на обновление m-го массива данных. Значение данной вероятности зависит от того, какая используется дисциплина обработки запроса, и опреде-

Таблица вероятностных временных и стоимостных характеристик дисциплин обработки запросов

Дисциплина	Xapa	ктеристика								
1	$\rho_j = \sum_{k=1}^N r_{jk} P_k(x_k) r_{kj} \psi_{jk}$	$Z = \sum_{j=1}^{N} \sum_{k=1}^{N} (U_{j} \psi_{jk} + V_{j} y_{k}) Z_{jk}(x_{k})$								
	$t_j = 2t_3(1 - \psi_{jj}) + \sum_{k=1}^{N} E_k(x_k)\Psi_{jk}$	$ZP = \sum_{j=1}^{N} \sum_{k=1}^{N} (U_{j} \psi_{jk} + V_{j} y_{k}) ZP_{jk}(x_{k})$								
2	$\rho_{j} = 1 - \prod_{k=1}^{N} [1 - r_{jk} P_{k}(x_{k}) r_{kj} \psi_{jk}]$	$Z = \sum_{j=1}^{N} \sum_{k=1}^{N} (U_{j} \psi_{jk} + V_{j} y_{k}) Z_{jk}(x_{k})$								
	$t_{j} = \max \left\{ 2t_{3} + \max_{k \in N_{j}; k \neq j} E_{k}(x_{k}); E_{j}(x_{kj}) \psi_{jj} \right\}$	$ZP = \sum_{j=1}^{N} \sum_{k=1}^{N} (U_{j} \psi_{jk} + V_{j} y_{k}) ZP_{jk}(x_{k})$								
3	$\rho_{j} = \sum_{n=1}^{K} r_{jj_{n}} P_{j_{n}}(x_{j_{n}}) r_{j_{n}j} \prod_{i=2}^{n} (1 - P_{j_{i-1}}(x_{j_{i-1}}) r_{j_{i-1}j})$	$Z = \sum_{j=1}^{N} \sum_{k=1}^{N} (U_j + V_j y_k) Z_{jk}(x_k)$								
	$t_j = (K+1)t_3 + \sum_{k=1}^N E_k(x_k)\psi_{jk} - B; \ B = t_3$ при $n(j,j) = 1$ или $n(j,j) = K$ и 0 в остальных случаях	$ZP = \sum_{j=1}^{N} \sum_{k=1}^{N} (U_j + V_j y_k) ZP_{jk}(x_k)$								
4	$\rho_{j} = 1 - \prod_{k=1}^{N} [1 - r_{jk} P_{k}(x_{k}) r_{kj} \psi_{jk}]$	$Z = \sum_{j=1}^{N} \sum_{k=1}^{N} \left\{ U_{j} \prod_{i=1}^{n(j,k)-1} [1 - r_{jj_{i}} P_{j_{i}}(x_{j_{i}}) r_{j_{i}j}] + V_{j} y_{k} \right\} Z_{jk}(x_{k}),$								
	$t_{j} = 2t_{3}(K - \psi_{jj}) + \sum_{k=1}^{N} E_{k}(x_{k})\psi_{jk}$	$j_{i} \in N_{j}; i = \overline{1, K}$ $ZP = \sum_{j=1}^{N} \sum_{k=1}^{N} \left\{ U_{j} \prod_{i=1}^{n(j,k)-1} [1 - r_{jj_{i}} P_{j_{i}}(x_{j_{i}}) r_{j_{i}j}] + V_{j} y_{k} \right\} ZP_{jk}(x_{k}),$								
2.5		$j_i \in N_j; i = \overline{1, K}$								
	Обозначения: $Z_{jk}(x_k) = 2(1-r_{jk})D_{jk} + [1-P_k(x_k)]E_k(x_k)h_k; \ ZP_{jk}(x_k) = 2D_{jk} + E_k(x_k)h_k; \ U_{ik} = \sum_{j=1}^n \psi_{jk}\psi_{jk}U_{ik}^{(j)}$ при $n(j,k) = n(j,i) + 1$ и 0 в остальных случаях; $U_{ik}^{(j)} = U_i^{n(j,i)-1} r_{j_{\alpha-1}j_{\alpha}}[1-P_{j_{\alpha}}(x_{j_{\alpha}})], j_{\alpha} = j, j_{\alpha} \in N_j$									
n(j, k) = n(j,	$i)+1$ и 0 в остальных случаях; $U_{ik}^{\circ\prime}=U_{i}$ \prod $r_{j_{o}}$	$\sum_{\alpha=1}^{j} j_{\alpha} [1 - P_{j_{\alpha}}(x_{j_{\alpha}})], j_{\alpha} = j, j_{\alpha} \in N_{j}$								



ляется по соответствующей формуле, приведенной в таблице [7]. Предположим, что значение этой вероятности равно вероятности события, заключающегося в обновлении по крайней мере одного из массивов данных, размещенного в одном из узлов сети с резервом, и последующего получения сообщения об успешном обновлении этого массива. С учетом сделанного предположения вероятность успешной обработки запроса на обновление массива данных

$$\beta_{nm}(X) = 1 - \prod_{i=1}^{N} [1 - r_{ni} P_i(x_{im}) r_{in}],$$

где $P_i(x_{im})$ — вероятность успешной обработки запроса в узле i, в котором размещен резерв объемом x_{im} ; r_{ni} — надежность канала связи между n-м и i-м узлом ВС (будем считать, что $r_{ni} = r_{in}$).

Обозначим через $T_{nj}(X)$ суммарное время ожидания ответов на все запросы, выданные при решении j-й задачи в n-м узле BC. Оно определяется по формуле

$$T_{nj}(X) = \sum_{m=1}^{M} \{ U_{jm} t'_{nm}(X) + U_{jm} t''_{nm}(X) \},$$

где $t'_{nm}(X)$ — время получения ответа на запрос к m-му массиву, выданный в n-м узле (его значение вычисляется по соответствующей формуле из таблицы); $t''_{nm}(X)$ — время обработки одного запроса на модификацию данных в m-м массиве, который был выдан в узле n. Значение данного времени также определяется по формуле из таблицы для дисциплины 2 обработки запросов при $N_i = \{i/x_{im} \neq 0\}$).

В таблице приведены формулы для расчета характеристик дисциплин обработки запросов: ρ_j — вероятности получения ответа на запрос, выданный в j-м узле BC; t_j — времени обработки запроса, выданного в узле j; ZP — размера средних производительных затрат на обработку запросов; Z — размера средних непроизводительных затрат на обработку запросов.

Стоимость S(X) функционирования РСОД в течение рассматриваемого нами периода времени будет складываться из стоимости решения всех задач системы и стоимости хранения резерва в узлах сети и, таким образом,

$$S(X) = \sum_{m=1}^{M} \sum_{j=1}^{J} W_{j} \sum_{n=1}^{N} \varphi_{jn} \left(V_{jm} \sum_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) + U_{jm} \min_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) \right) + L_{m} \sum_{n=1}^{N} s_{n} x_{nm},$$

где $ZP_{ni}(x_{im}) = 2D_{ni} + E_i(x_{im})h_i$, s_n — стоимость хранения одного бита информации в n-м узле в течение рассматриваемого периода времени.

Сформулируем задачу оптимального оперативного резервирования M взаимосвязанных массивов данных при использовании в качестве критерия минимума стоимостных затрат на функционирование системы S(X).

Необходимо найти такое оптимальное размещение резерва используемых в РСОД массивов данных по узлам ВС, которое обеспечит минимальные стоимостные затраты на функционирование системы. Найденное решение задачи должно соответствовать ограничениям на объем резерва, размещаемого в каждом из узлов ВС, и должно обеспечивать затраты времени на ожидание ответов на запросы не более заданного лимита.

Таким образом, найти

$$S(X) \rightarrow \min$$

при ограничениях:

$$P_{nj}(X) \ge \overline{P}_j, \quad T_{nj}(X) \le \overline{T}_j, \quad \sum_{m=1}^{M} x_{nm} L_m \le Q_n,$$

$$j = \overline{1, J}; \quad n = \overline{1, N}.$$

Аналогично формулируются задачи поиска оптимального размещения в ВС резерва взаимосвязанных массивов, используемых в РСОД, при других критериях оптимальности решения. Например, при использовании такого критерия оптимальности размещения резерва, как максимум вероятности успешного решения всех задач системы за заданный период времени, задача формулируется следующим образом:

$$P(X) = \prod_{j=1}^{J} \prod_{n=1}^{N} [P_{nj}(X)]^{W_j} \to \max$$

при ограничениях

$$S(X) \le \overline{S}, \quad T_{nj}(X) \le \overline{T}_j, \quad \sum_{m=1}^{M} x_{nm} L_m \le Q_n,$$

$$j = \overline{1, J}, \quad n = \overline{1, N}.$$

4. РЕЗЕРВИРОВАНИЕ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ БАЗАХ ДАННЫХ

Рассмотрим ситуацию, когда необходимо найти оптимальное размещение резерва взаимосвязанных массивов для распределенной базы данных (РБД), в адрес которой пользователями ВС формируются поток информационных запросов и поток запросов на модификацию данных.

Обозначим: J — множество номеров информационных запросов к массивам, обрабатываемым средствами РБД; $J^{\rm o}$ — множество номеров различных запросов на модификацию данных в массивах РБД; M_j — множество номеров массивов, используемых при обработке j-го информационного за-



проса, $j \in J$; $M_j^{\rm o}$ — множество номеров массивов, которые модифицируются при выполнении одного запроса $j^{\rm o} \in J^{\rm o}$. В каждый фиксированный интервал времени в адрес РБД в n-м узле сети в среднем формируется U_{nj} запросов j-го типа (информационный запрос к массиву данных) и $V_{nj}^{\rm o}$ запросов $j^{\rm o}$ -го типа (модификация массива данных).

Вероятность $R_n(X)$ получения ответов на все запросы, выданные в n-м узле за рассматриваемый интервал времени (т. е. надежность обслуживания пользователей n-го узла сети, генерирующих поток запросов к РБД) определяется по формуле:

$$R_{n}(X) =$$

$$= \prod_{j \in J} \left[\prod_{m \in M_{j}} \rho_{nm}(X) \right]^{U_{nj}} \prod_{j^{\circ} \in J^{\circ}} \left[\prod_{m \in M_{i}^{\circ}} \beta_{nm}(X) \right]^{V_{nj^{\circ}}}.$$

Всеми пользователями n-го узла BC, работающими с РБД, за рассматриваемый интервал времени будет затрачено $T_n(X)$ единиц времени на ожидание ответов на выданные ими запросы:

$$T_{n}(X) = \sum_{j \in J} U_{nj} \sum_{m \in M_{j}} t'_{nm}(X) + \sum_{j^{\circ} \in J^{\circ}} V_{nj^{\circ}} \sum_{m \in M^{\circ}} t''_{nm}(X).$$

Стоимость функционирования РБД за рассматриваемый интервал времени (затраты на хранение резерва и обработку запросов)

$$S(X) = \sum_{n=1}^{N} \left[\sum_{j \in J} U_{nj} \sum_{m \in M_{j}} \min_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) + \right] + \sum_{j^{\circ} \in J^{\circ}} V_{nj^{\circ}} \sum_{m \in M_{j}^{\circ}} \sum_{i/x_{im} \neq 0} ZP_{ni}(x_{im}) + \sum_{m=1}^{M} L_{m} \sum_{n=1}^{N} S_{n}x_{nm}.$$

Полученные формулы позволяют сформулировать задачу оптимизации резервирования массивов данных в РБД по критерию максимума надежности обслуживания пользователей базы данных — необходимо обеспечить максимальную вероятность обработки всех запросов пользователей к РБД, сформированных за рассматриваемый интервал времени, при ограничениях на объем резерва, размещаемого в каждом из узлов сети, стоимость функционирования РБД, общее время ожидания ответов на запросы, выданные пользователями одного узла за рассматриваемый интервал времени:

$$R(X) = \prod_{n=1}^{N} R_n(X) \to \max$$

при ограничениях
$$T_n(X) \leq \overline{T}_n$$
, $S(X) \leq \overline{S}$,
$$\sum_{m=1}^M x_{nm} L_m \leq Q_n, \ n=\overline{1,N}.$$

Аналогично формулируются и задачи оптимального резервирования массивов данных в РБД с другими критериями оптимизации: минимум стоимости функционирования РБД за заданный интервал времени; минимум суммарного времени ожидания ответов на запросы пользователей РБД, сформированные ими за заданный интервал времени.

5. РАЗМЕЩЕНИЕ РЕЗЕРВА ИНФОРМАЦИИ С УЧЕТОМ ИНТЕНСИВНОСТИ ТРАФИКА ЗАПРОСОВ

Оптимальное распределение оперативного резерва массивов данных по узлам ВС находится в непосредственной зависимости от трафика запросов к резервируемым массивам, возникающих в узлах сети при решении задач и формируемых пользователями сети. Выполним качественный анализ влияния трафика запросов, возникающих в узлах ВС, на оптимальное распределение оперативного резерва массивов данных. Анализ выполним на примере однородной полносвязной ВС в смысле размера стоимостных затрат на хранение резерва и обработку запросов — S(Y). Будем считать, что запросы на модификацию массива данных адресуются одновременно во все узлы сети с резервом, а информационные запросы к массиву обрабатываются в соответствии с дисциплиной 1. Затраты будем вычислять по формуле

$$S(Y) = [h + 2D + h^*(T_Y - T + 2t_3)] \sum_{j=1}^{N} (U_j + KV_j) + sK - (2D + h^*2t_3) \sum_{j=Y} (U_j + V_j),$$
 (13)

где D — стоимость передачи запроса/ответа по каналам связи BC; h — стоимость обработки одного запроса в узле с резервом; s — стоимость хранения резерва в одном из узлов; Y — множество индексов узлов сети, по которым распределен резерв (распределение резерва по BC); K = |Y|; T_Y — среднее время обработки запроса в одном из узлов сети при распределении резерва Y; T — заданная верхняя граница времени получения ответа на запрос; h^* — стоимость (штраф) превышения временем получения ответа величины T.

Сделаем предположения:

- информационные запросы, возникающие в узлах без резерва, равномерно распределяются между всеми K узлами сети с резервом;
- запросы, поступающие в каждый из узлов с резервом, образуют пуассоновский поток с пара-



метром λ_y ; время обработки запросов распределено по показательному закону;

— для каждого узла сети $j = \overline{1, N}$ выполняются равенства: $U_i = U; \ V_j = V.$

Основываясь на сделанных предположениях, получим, что $T_{\gamma} = t/(1-\lambda_{\gamma}t)$, где t — среднее время обработки запроса, а $\lambda_{\gamma} = NV + NK^{-1}U$.

Предположим, что размещение резерва Y мы получили на основе распределения Y путем размещения в узле i дополнительного резерва. Определим, при каких соотношениях между значениями U и V справедливо неравенство

$$S(Y') > S(Y), \tag{14}$$

в котором $Y' = Y \cup \{i\}; (K' = |Y'| = K + 1).$

При сделанных предположениях из формулы (13) следует, что

$$S(Y') = S(Y) - N(U + KV)(T_Y - T_Y)h^* + S - (U + V)(2D + h^*2t_3) + + NV[h + 2D + h^*(T_Y - T) + h^*2t_2],$$
(15)

а неравенство (14) будет эквивалентно неравенству

$$V[Nh + 2D(N-1) + Nh^{*}(T_{Y'} - T) + + h^{*}2t_{3}(N-1) - NK(T_{Y} - T_{Y})h^{*}] > > U[Nh^{*}(T_{Y} - T_{Y}) + 2D + h^{*}2t_{3}] - S.$$
 (16)

Достаточным условием выполнения неравенства (16) служит соотношение между значениями U и V:

$$V > U[Nh^*(T_Y - T_Y) + 2D + h^*2t_3] \times \times [Nh + 2D(N - 1) + Nh^*(T_Y - T) + + h^*2t_3(N - 1) - NK(T_Y - T_Y)h^*]^{-1}.$$
(17)

Рассмотрим частные случаи.

1. Значение S(Y) зависит главным образом от затрат на использование каналов связи, т. е. $D\gg h^*$, $D\gg h$. В этом случае соотношение (17) приводится к виду

$$V \ge U(N-1)^{-1}. (18)$$

2. Превышение временем получения ответа на запрос величины T несущественно, т. е. $h^* \ll h$, $h^* \ll D$.

Соотношение (17) выполняется при условии

$$V \ge U(N-1)(h/2D+1)]^{-1}$$
. (19)

3. Учитывается главным образом время получения ответа на запрос, т. е. $h^* \gg D$, $h^* \gg h$. В этом случае выражение (17) эквивалентно соотношению

$$V \ge U[N(T_Y - T_Y) + 2t_3][N(T_Y - T) + (N-1)2t_2 - N_\nu(T_V - T_\nu)]^{-1}.$$
 (20)

Полученные качественные результаты (при сделанных предположениях) сформулируем в виде утверждения,

Утверждение 1. В однородной полносвязной вычислительной сети при выполнении соотношения (17) (или соотношений (18)—(20)) соответственно для рассмотренных случаев 1—3 наименьшие стоимостные затраты на хранение резерва и на обработку запросов достигаются при размещении резерва в одном узле сети.

Из выражения (15) следует, что неравенство $S(Y') \leq S(Y)$ справедливо при выполнении соотношения

$$U > V\{Nh + 2D(N-1) + h^*[N(T_Y - T) - 2t_3(N-1) - NK(T_Y - T_Y)]\}B^{-1} + SB^{-1}, (21)$$

где
$$B = Nh^*(T_V - T_V) + 2D + h^*2t_3$$
.

Рассмотрим частные случаи:

1) если $D\gg h^*,\ D\gg h,\ D\gg S,$ то соотношение (21) эквивалентно неравенству

$$U > V(N-1); \tag{22}$$

2) при $h^* \ll h, h^* \ll D$ и $h^* \ll S$ соотношение (21) приводится к виду

$$U > V(N-1) + (S + NhV)(2D)^{-1};$$
 (23)

3) для $h^* \gg D$ и $h^* \gg h$ получим:

$$U > V[N(T_{Y'} - T) + 2t_{3}(N - 1) - NK(T_{Y'} - T_{Y'})][N(T_{Y'} - T_{Y'}) + 2t_{3}]^{-1}.$$
 (24)

Обобщим полученные качественные результаты в виде утверждения, справедливого при сделанных предположениях.

Утверждение 2. В однородной полносвязной вычислительной сети при выполнении соотношения (21) (или соотношений (22)—(24)) соответственно для рассмотренных случаев 1—3 наименьшие стоимостные затраты на хранение резерва и обработку запросов обеспечивает размещение резерва во всех узлах, в которых возникают запросы.

Пример. Покажем на примере решения задачи оптимального оперативного резервирования массивов данных в ВС то, как влияет на оптимальное размещение резерва изменение значений параметров самой сети. Рассмотрим результат решения задачи оптимального резервирования на примере сегмента однородной полносвязной ВС из пяти узлов. Параметры узлов рассматриваемого сегмента сети (в силу однородности сети индексы узлов опущены):

$$\theta = 1.5;$$
 $\tau = 0.5;$ $U = 20;$ $V = 7;$ $q = 0.01;$ $S = 2.0;$ $r^2 = 0.985;$ $2D = 9.0;$ $h = 3.0;$ $N = 5.$

В качестве критерия задачи примем минимум стоимостных затрат на обработку запросов и хранение резерва.



Для рассматриваемого нами примера сегмента сети с указанными параметрами выполняется соотношение (18). Следовательно, справедливо и утверждение 1, из которого следует, что при заданных параметрах сети резерв целесообразно разместить только в одном из узлов.

Проверка данного утверждения путем решения задачи оптимального оперативного резервирования методом полного перебора показывает, что для рассматриваемого фрагмента сети оптимальными являются пять равнозначных (в силу однородности сети) распределений резерва: (20000), (02000), (00200), (00020) и (00002). Эти распределения резерва обеспечивают одинаковое значение целевой функции S(Y) = 1792,07.

Если же в параметрах рассматриваемого сегмента сети произойдут изменения: V=3, h=1,5, то в этом случае будет выполняться соотношение (23) и будет уже справедливо утверждение 2. В соответствии с этим утверждением при измененных параметрах ВС наиболее целесообразно размещение резерва в каждом из узлов сегмента, в котором возникают запросы. Решение задачи оптимального оперативного резервирования методом полного перебора при новых параметрах фрагмента сети доказывает справедливость утверждения 2 и дает нам одно оптимальное распределение оперативного резерва — (11111), которое обеспечивает следующее значение целевой функции S(Y)=1337,49.

ЗАКЛЮЧЕНИЕ

Рассмотрены три основные стратегии оперативного резервирования массивов постоянных и текущих оперативных данных и указаны особенности применения этих стратегий в системах обработки данных, работающих в распределенных вычислительных сетях. Описаны различные варианты хранения резерва в вычислительной сети и несколько возможных дисциплин обработки информационных запросов и запросов на модификацию данных массивов. Сформулирована в общем виде задача распределения информационного резерва по узлам вычислительной сети, оптимального по критерию минимума суммарных затрат на обработку поступающих запросов и хранение резерва в узлах сети. Также сформулированы задачи оптимального резервирования массивов данных в вычислительных сетях по критериям максимума вероятности успешной обработки запроса к распределенным системам обработки данных и минимума среднего времени получения ответа на запрос. Даны формулировки задач оптимального оперативного резервирования массивов данных в вычислительных сетях с учетом различного рода взаимосвязей между резервируемыми массивами и задач оптимального резервирования в распределенных базах данных. Приведены методы решения поставленных задач оптимизации. На примере однородной полносвязной сети ЭВМ выполнен качественный анализ влияния интенсивностей запросов, возникающих в отдельных узлах сети, на распределение оперативного резерва массивов данных. Изложенные в статье результаты имеют практическую ценность и могут быть использованы при проектировании распределенных систем обработки данных в целях повышения безопасности их функционирования.

ЛИТЕРАТУРА

- NASA publication SP-480, Far Travelers -The Exploring Machines, Oran W. Nicks, 1985. URL: http://history.nasa.gov/SP-480/ch3.htm (дата обращения 14.07.2016).
- NASA Space Science Data Coordinated Archive. Mariner 1. NSSDCA ID: MARIN1. — URL: http://nssdc.gsfc.nasa.gov/ nmc/spacecraftDisplay.do?id=MARIN1 (дата обращения 23.08.2016).
- Mariner Venus Final Project Report (NASA SP-59, 1965). —
 P. 87. URL: http://history.nasa.gov/SP-4407/ETUv5.pdf (дата обращения 14.07.2016).
- 4. *History* of the OCLC Research Library Partnership. URL: http://www.oclc.org/research/partnership/history.html обращения 14.07.2016).
- 5. *Информационная* безопасность систем организационного управления. Теоретические основы: в 2 т. / Н.А. Кузнецов, В.В. Кульба, Е.А. Микрин и др.; отв.ред. Н.А. Кузнецов, В.В. Кульба; Ин-т проблем передачи информации РАН. М.: Наука, 2006.
- 6. *Кульба В.В., Мамиконов А.Г., Шелков А.Б.* Резервирование программных модулей и информационных массивов в ACV // Автоматика и телемеханика. 1980. № 8. С. 133—141.
- 7. *Кульба В.В., Сомов С.К., Шелков А.Б.* Резервирование данных в сетях ЭВМ. Казань: Изд-во Казанского гос. ун-та, 1987. 175 с.
- 8. *Сомов С.К.* Резервирование программных модулей и информационных массивов в сетях ЭВМ: дисс. канд. техн. наук. М.: ИПУ РАН, 1983. 217 с.
- Гурин Л.С., Дымарский Я.С., Меркулов А.Д. Задачи и методы оптимального распределения ресурсов. — М.: Сов. радио, 1968. — 463 с.
- Machmoud S., Riordon J.S. Optimal Allocation of Resources in Distributed Information networks // ACM Trans. on Database Systems. — 1976. — Vol. 1, N 4. — P. 66—78.
- Chu W.W. File Allocation in an Multiple Computer System // IEEE Trans. on Computers. — 1969. — Vol. C-18, N 10. — P. 885—889.
- Casey R.G. Allocations of copies of a file in an Information Network // AFIPS Conference Proceedings. — 1972. — Vol. 40. — P. 617—625.
- 13. *Akoka J.* Design of optimal distributed database systems. Distributed Data Bases // Proc. Int. Symp. 1980. P. 229—246.
- Ramamoorthy C.V., Wah B.W. Data management in distributed data bases // AFIPS Conference Proceedings. 1978. Vol. 48. P. 667—680.
- 15. *Клейнрок Л*. Вычислительные системы с очередями. М.: Мир, 1979. 600 с.

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

Микрин Евгений Анатольевич — академик РАН, ген. конструктор, ПАО «Ракетно-космическая корпорация «Энергия» им. С.П. Королева», г. Королев, ⊠ Eugeny.Mikrin@rsce.ru,

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

УДК 651/689

МАТЕМАТИЧЕСКИЕ МОДЕЛИ И АЛГОРИТМЫ ПОСТРОЕНИЯ ДОПУСТИМЫХ И ОПТИМАЛЬНЫХ МАРШРУТОВ ДВИЖЕНИЯ И ДОСТАВКИ ГРУЗОВ

Ю.А. Зак

Предложены постановки и установлены математические свойства задач построения допустимых и оптимальных маршрутов выполнения заданий одним или несколькими видами транспортных средств в условиях заданной системы ограничений на сроки в пунктах приема и доставки грузов и невозможности одновременного выполнения двух или нескольких заданий. На основе установленных свойств допустимых и оптимальных последовательностей разработаны алгоритмы решения задач модифицированными последовательными методами оптимизации, которые проиллюстрированы числовыми примерами. Приведены результаты вычислительных экспериментов.

Ключевые слова: последовательности выполнения заданий, допустимые и оптимальные маршруты, оценки частичных планов, модифицированные последовательные алгоритмы оптимизации.

ВВЕДЕНИЕ

Построение эффективных маршрутов движения транспортных средств, обеспечивающих с наименьшими затратами времени и средств доставку грузов в установленных объемах и в определенные договорными обязательствами сроки, — одна из важнейших задач транспортной логистики [1, 2]. Оптимальным считается маршрут, по которому возможно доставить пассажиров или груз в кратчайшие сроки и предусмотренные графиками поставки сроки с минимальными затратами времени или стоимости.

Рассматриваемая в работе проблема относится к классу задач Vehicle Routing Problems with Pickup and Delivery (VRPD), в которых рассматривается осуществление адресных доставок одного или нескольких видов грузов некоторому числу потребителей с пунктов производства, хранения или распределения грузов. Средства доставки транспортные средства — (TC) ограниченной грузоподъемности могут состоять из одного или нескольких видов машин с различными техническими ха-

рактеристиками. После выполнения заданий все ТС должны вернуться в свой начальный пункт базирования. В частных случаях рассматриваются дополнительные проблемы, связанные с необходимостью возвращения остатков грузов от потребителей в пункты их получения (депот). В задачах Vehicle Routing Problems with Time Windows (VRPTW) предусмотрены дополнительные ограничения на временные окна приема грузов для некоторого числа или всех потребителей. В самом общем случае эти задачи могут иметь несколько целей, как например, выполнение всех заданий в кратчайшие сроки, минимум используемых и необходимых ТС, минимизация общего времени работы всех ТС в пути, минимизация суммарного времени ожидания обслуживания всех ТС (из-за наличия временных окон), а также суммарного расстояния поездки и др. В литературе описаны различные постановки задач данного класса (см., например, работы [3—21]). Основное внимание при этом уделялось постановкам задач, в которых груз в установленном объеме, не превышающем грузоподъемности транспортного средства (ТС), из



одного и того же пункта (депот) одной или несколькими идентичными машинами должен быть доставлен некоторому множеству потребителей. В частных случаях рассматриваются задачи с одним местом приема грузов, одним ТС (которые сводятся к обобщенной задаче коммивояжера), ситуации, когда потребности каждого потребителя могут быть удовлетворены одной поставкой. Математическая модель задачи в такой постановке аналогична известной в литературе задаче Job-shop scheduling with sequence-dependent setup times [15].

Авторами рассматриваются математические модели сформулированных проблем в виде моделей целочисленного линейного программирования большого размера с двух- или трехиндексным обозначением переменных [9, 11, 17] и построения допустимых экстремальных путей на графе [7, 17, 21].

Задачи VRPD и VRPTW относятся к классу NP-сложных проблем. В качестве точных методов рассматривались методы динамического программирования и модификации метода ветвей и границ (Column Generation) с использованием препроцессоров и процедур декомпозиции [18, 20]. На практике большое число клиентов и сложная система ограничений не позволяют получить точное решение задачи. В связи с этим много публикаций посвящено разработке эвристических алгоритмов решения задачи в комбинации с методами разбиения множества пунктов на кластеры, применением методов декомпозиции, случайного поиска и генетических алгоритмов [8, 12, 18]. Определенный интерес представляют следующие подходы: маршрут образующие и улучшающие маршрут процедуры, Savings-Heuristik [12], PSO — алгоритм оптимизации роя частиц [13], алгоритм ASC [14], MCPSO [11] и др.), генетические алгоритмы [16], а также алгоритмы направленного случайного поиска (Tabu-Search [10], поиск с запретами [19]), реализация одностороннего ветвления в методе ветвей и границ. При решении задач применялись препроцессоры, позволяющие упростить граф взаимосвязей между пунктами, исключив заведомо неэффективные и недопустимые в смысле выполнения ограничений связи, а также определить наиболее ранние и допустимые самые поздние времена выезда из пунктов [14, 17], а также методы кластерного анализа [20]. Главной проблемой при этом было получение допустимого решения. При этом не приводились оценки точности полученного решения. Эффективность предлагаемых подходов подтверждалась результатами вычислительного эксперимента.

В отличие от известных публикаций, в данной работе предлагается другая постановка задачи, предусматривающая такие особенности:

- груз одним или несколькими одинаковыми или различными по грузоподъемности и техническим характеристикам типами ТС может доставляться каждому из множества потребителей из различных пунктов приема груза; начальный пункт может не совпадать с местом приема грузов;
- большие объемы поставок при ограниченной грузоподъемности ТС могут требовать выполнения нескольких маршрутов между соответствующей парой пунктов;
- одновременная загрузка TC грузами для различных потребителей недопустима;
- учитываются затраты на погрузку и разгрузку груза;
- временные окна могут быть заданы не только для доставки, но и для приема грузов.

Применение точных, а также эвристических и приближенных методов для решения задач этого класса и исследование свойств допустимых и оптимальных маршрутов позволили разработать алгоритмы эффективного отсева недопустимых и заведомо неоптимальных маршрутов. Полученные оценки позволяют оценить точность приближенных решений в условиях решения задач большой размерности, а также установить факт несовместности системы ограничений на ранних этапах решения.

1. ПОСТАНОВКА ЗАДАЧИ И МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

В настоящей статье рассматриваются постановки задачи построения допустимых и оптимальных маршрутов движения ТС в условиях, когда задано некоторое подмножество пунктов $ilde{J}^1$ приема и подмножество пунктов $ilde{m{J}}^2$, куда необходимо доставить груз. Транспортное средство должно выполнить некоторое множество заданий $k \in K = \{1, 2, ..., K\}$. Задано множество пар пунктов приема и доставки грузов для каждого задания $\tilde{\Pi} = \{\pi_1, \, ..., \, \pi_k, \, ..., \, \pi_K \},$ где $\pi_k = (i_k, j_k), \, i_k \in \, \tilde{\boldsymbol{J}}^1, j_k \in \, \tilde{\boldsymbol{J}}^2;$ а также времена и (или) стоимости переезда ТС между каждой парой пунктов транспортной сети. Каждое из заданий заключается в том, что ТС должно взять груз в заданном пункте $i \in \tilde{J}^1$ и доставить его в некоторый фиксированный пункт из подмножества $j \in \tilde{J}^2$. Грузоподъемность TC позволяет выполнить каждое из заданий множества \tilde{K} . Одновременное выполнение двух или нескольких заданий недопустимо. Условиями задачи могут предусматриваться



потери времени или затраты на погрузку и разгрузку грузов, а также ограничения на сроки доставки и загрузки груза в подмножествах пунктов \tilde{J}^1 и (или) \tilde{J}^2 . Это множество заданий может выполняться одним или несколькими видами одинаковых или различных по техническим характеристикам TC. Необходимо найти маршруты доставки грузов каждого TC, обеспечивающие выполнение всех ограничений задачи и минимизирующие суммарные затраты средств на выполнение всех заданий.

Автору не известны постановки и методы решения рассматриваемого класса задач в условиях заданной системы условий и ограничений. В отличие от известных задач коммивояжера [22] и *N* коммивояжеров [23], в общем случае, не существует решения задачи в виде замкнутых гамильтоновых маршрутов, так как ТС придется заезжать в некоторые пункты более чем по одному разу. Также в некоторых случаях необходимы простои ТС до нижней границы допустимого времени обслуживания пункта.

В данной статье исследуются математические особенности рассматриваемых задач этого класса, строятся оценки оптимальных и возможности получения допустимых решений, формулируются правила предпочтения и определения перспективности анализируемых вариантов. Показано, что для каждого анализируемого частичного плана существует только некоторое допустимое подмножество возможных продолжений, которое должно быть выделено с целью сокращения множества анализируемых вариантов. Все это потребовало разработки оригинальных методов вычисления оценок, стратегий ветвления, а также (на основе «оптимистических» оценок возможности выполнения ограничений) разработать правила отсева недопустимых продолжений. На этой основе предлагаются эффективные методы решения рассматриваемых задач с помощью последовательных алгоритмов оптимизации (модифицированные методы ветвей и границ и динамического программирования), учитывающие специфику каждой конкретной задачи. Приведенные алгоритмы решения иллюстрируются числовыми примерами. Показано, что в ряде случаев в предлагаемых автором алгоритмах наличие жестких ограничений на время доставки и приема грузов в различных пунктах позволяет уже на ранних этапах исключить маршруты, не содержащие допустимых решений, что существенно сокращает число анализируемых вариантов.

Транспортное средство должно доставить грузы от некоторого подмножества производителей и

складов дистрибьютеров \tilde{J}^1 подмножеству потребителей ${ ilde J}^2$. Общее число пунктов загрузки и разгрузки ТС, включая начальный пункт (депот) равно $(n+1)-\tilde{I}=(\tilde{J}^1\cup \tilde{J}^2)=\{i=0,1,...,n\}$. Каждое из заданий заключается в том, что TC должно взять груз в заданном пункте $i \in \tilde{J}^1$ и доставить его в некоторый фиксированный пункт из подмножества $j \in \tilde{J}^2$. Необходимо выполнить некоторое множество заданий $k \in \tilde{K} = \{1, 2, ..., K\}$. Если грузоподъемность ТС не позволяет выполнить какое-либо из рассматриваемых заданий, то последнее представляется в виде нескольких, объемы поставок которых могут быть различными. Одновременное выполнение двух заданий, т. е. загрузка ТС двумя видами грузов, недопустима. Следовательно, все задания должны выполняться последовательно друг за другом. В начальный момент времени TC находится на базе (i = 0), на которую должно вернуться при завершении выполнения всех перевозок. Грузы могут перевозиться одним или несколькими одинаковых или различных по техническим характеристикам ТС. Могут быть заданы различного вида ограничения на времена выполнения отдельных заданий.

Пусть заданы пары пунктов погрузки и доставки грузов для каждого из K заданий: $G(k)=(i_k,j_k)$, где i_k — индекс пункта, в котором принимается груз, а j_k — индекс пункта, куда он должен быть доставлен; τ_k^+ и τ_k^- — время, необходимое для загрузки товара в пункте k и выгрузки груза в пункте j_k при выполнении k-го задания, $\bar{\tau}_k = \tau_k^+ + \tau_k^-$ — суммарные потери на погрузку и выгрузку грузов при выполнении k-го задания; $C = \|c_{ij}\|$ и $T = \|t_{ij}\|$, i, j = 0, 1, ..., n — соответственно матрицы стоимости и времени переезда TC между пунктами.

Кроме того, для некоторого подмножества заданий могут быть заданы: $H^1(k) \in [h^1_{ki}, h^2_{ki}],$ $H^2(k) \in [h^1_{kj}, h^2_{kj}]$ — соответственно промежутки времени, в которые необходимо принять груз в пункте i_k и доставить его в пункт j_k .

Рассматриваются следующие задачи транспортной логистики.

Задача 1. План перевозок осуществляется одним ТС. Не существует никаких ограничений на времена прибытия в пункты приема и доставки грузов. Необходимо найти маршрут движения ТС, обеспечивающий минимальные затраты времени или стоимости выполнения всех заданий.



Задача 2. При выполнении плана перевозок одним ТС заданы различного вида ограничения на сроки выполнения отдельных заданий: граничные сроки завершения выполнения, временные окна приема или доставки грузов в некоторое подмножество пунктов и др. Необходимо найти расписание выполнения всех заданий, обеспечивающее минимальные затраты времени или стоимости выполнения всех заданий.

Задача 3. Все задания выполняются *М* одинаковыми по техническим характеристикам TC. Ограничения на времена выполнения отдельных заданий отсутствуют. Необходимо определить непересекающиеся подмножества заданий, выполняемых каждым из TC, а также маршруты движения каждого из них, обеспечивающие минимальные затраты времени или стоимости выполнения всех заданий.

Задача 4. Все задания выполняются M различными по техническим характеристикам ТС. Заданы τ_k^+ и τ_k^- , $\bar{\tau}_k = \tau_k^+ + \tau_k^-$, а также $C^m = \|c_{ij}^m\|$ и $T^m = \|t_{ij}^m\|$, i,j=0,1,...,n, для каждого типа ТС, m=1,...,M. В условиях, когда заданы различного вида ограничения на сроки выполнения отдельных заданий и времена работы каждого из ТС, необходимо определить непересекающиеся подмножества заданий, выполняемых на каждом из ТС, а также маршруты движения каждого из них, обеспечивающие минимальные затраты времени или стоимости выполнения всех заданий. \bullet

Определим время, необходимое ТС для завершения выполнения k-го задания при переезде из пункта $j_{k-1} \in \tilde{J}^2$ приема груза после выполнения стоящего в последовательности перед ним (k-1)-го задания:

$$q_{k-1,k}^{1} = \bar{\tau}_{k} + t(j_{k-1}, i_{k}) + t(i_{k}, j_{k}), \qquad (1)$$

где $t(i_k,j_k)$ — время движения ТС из i_k -го пункта приема до j_k -го пункта доставки груза k-го задания, $t(j_{k-1},i_k)$ — время движения ТС из j_{k-1} -го пункта доставки (k-1)-го задания до i_k -го пункта приема груза k-го задания.

Время, необходимое ТС для прибытия в i_k -й пункт приема груза при выполнении k-го задания после прибытия в i_{k-1} -й пункт, и время до выполнения (k-1)-го задания соответственно

$$q_{k-1,k}^2 = t(j_{k-1}, i_k), \ q_{k-1,k}^3 =$$

$$= \bar{\tau}_{k-1} + t(i_{k-1}, j_{k-1}) + t(j_{k-1}, i_k). \tag{2}$$

Пусть на некотором s-м этапе решения задачи выполнено подмножество заданий \tilde{K}^s и построена подпоследовательность \tilde{W}^s доставки грузов в подмножество пунктов $\tilde{J}^{2,\,s}$, которую назовем s-м частичным планом. Обозначим: $\hat{J}^{2,\,s}$ — подмножество оставшихся пунктов доставки грузов в s-м частичном плане; $\hat{K}^{s,\,-}$ — подмножество невыполненных заданий; (k-1) — индекс последнего задания в подпоследовательности $\tilde{J}^{2,\,s}$; $F(\tilde{J}^{2,\,s})$ — функция цели этого частичного плана (стоимость или время завершения выполнения этих заданий).

Нижние границы времен завершения выполнения следующего k-го, $k \in \hat{J}^{2,s}$, стоящего первым после этой подпоследовательности $\tilde{J}^{2,s}$, задания:

— на начальном этапе решения

$$\xi(q_{0,k}^1) = \min_{k \in \tilde{K}} [\bar{\tau}_k + t_{ij}(k) + t_{ji}(0,k)],$$

$$\xi(q_{k,l}^{1}) = \min_{k \in \tilde{K}} [\bar{\tau}_{k} + t_{ij}(k) + t_{ji}(l, k)];$$

$$\xi(q_{0,k}^2) = \min_{k \in \tilde{K}} t_{ji}(0, k), \quad \xi(q_{k,l}^2) = \min_{k \in \tilde{K}} t_{ji}(l, k);$$

— на некотором *s*-м этапе решения некоторого $l \in \hat{K}^{s,-}$ -го подлежащего выполнению задания

$$\xi(q_{k,l}^{1}) = \min_{k \in \tilde{K}} [\bar{\tau}_{k} + t(i_{l}, j_{l}) + t(j_{k}, i_{l})],$$

$$\xi(q_{k,l}^{2}) = \min_{k \in \tilde{K}} (j_{k}, i_{l}).$$

Значения этих параметров будут использованы в дальнейшем для оценок возможности выполнения ограничений на допустимые ранние сроки прибытия в пункты назначения и поздние сроки завершения выполнения заданий. Отметим, что наличие ограничений на ранние сроки прибытия в некоторые пункты (значения h_{ki}^1 , h_{kj}^1) могут привести к некоторым потерям, связанными с простоями TC (θ_{ki} и θ_{kj}), которые должны быть учтены в выражениях (1) и (2).

Пусть последовательность выполнения заданий ТС имеет вид

$$\tilde{W} = \{w_0, w_1, ..., w_k, ..., w_K, w_0\}. \tag{3}$$

Время завершения первого стоящего в последовательности $ilde{W}$ задания определяется выражением

$$q_{0,k}^1 = \bar{\tau}_k + t(0,k) + t(i_k,j_k).$$



Время, необходимое после прибытия в пункт приема груза последнего стоящего в последовательности \tilde{W} задания k до прибытия в пункт 0-й базирования TC, определяется выражением

$$q_{k,0}^2 = t(i_k, 0), \quad q_{k,0}^3 = \bar{\tau}_k + t(i_k, 0) + t(i_k, j_k).$$

Суммарное время, необходимое для выполнения всех заданий, выполняемых в последовательности (3), определяется выражением

$$T(\tilde{W}) = q_{0, w_1}^1 + \sum_{\substack{(w_{9-1}, w_9) \in \{\tilde{W}/(w_0, w_1)\}}} q_{w_{9-1}, w_9}^1 + \sum_{k=1}^K \theta_k, \quad (4)$$

или

$$T(\tilde{W}) = q_{K,0}^2 + \frac{\sum_{(w_{9-1}, w_{9}) \in \{\tilde{W}/(w_{K}, w_{0})\}} q_{w_{9-1}, w_{9}}^1 + \sum_{k=1}^K \theta_k, \quad (5)$$

где $\theta_k = \theta_{ik} + \theta_{jk}$ и θ_k^m — суммарное время простоев из-за ожидания допустимого времени начала обслуживания при выполнении k-го задания на одном и на m-м TC; $\{\tilde{W}/(w_0, w_1)\}$, $\{\tilde{W}/(w_K, w_0)\}$ — соответственно множество пар заданий из последовательности (3) за исключением пар (w_0, w_1) и (w_K, w_0) .

В дальнейшем ограничимся рассмотрением только критериев минимизации времени выполнения всего множества заданий в виде (4) или (5).

Если грузы доставляются несколькими видами TC и определены маршруты движения каждого TC

$$\tilde{W}^{m} = \{w_{0}, w_{1}^{m}, ..., w_{k}^{m}, ..., w_{K^{m}}^{m}, w_{0}\},\$$

$$m = 1, ..., M.$$

где K^m — число заданий, выполняемых m-м TC, то критерий оптимальности расписания выполнения заданий может быть представлен в виде:

— минимизация суммарного времени выполнения всех заданий всеми TC:

$$F_{1}(\tilde{W}^{1}, ..., \tilde{W}^{m}, ..., \tilde{W}^{M}) = \sum_{m=1}^{M} \left\{ q_{0, w_{1}^{m}}^{1} + \sum_{(w_{0-1}^{m}, w_{0}^{m}) \in \{\tilde{W}^{m}/(0, w_{1}^{m})\}} q_{w_{9-1}^{m}, w_{9}^{m}}^{1} + \sum_{k=1}^{K^{m}} \theta_{k}^{m} \right\}, \quad (6)$$

$$F_{1}(\tilde{W}^{1}, ..., \tilde{W}^{m}, ..., \tilde{W}^{M}) = \sum_{m=1}^{M} \left\{ q_{W_{K}^{m}, 0}^{2} + \sum_{(w_{3-1}^{m}, w_{3}^{m}) \in \left\{ \tilde{W}^{m} / \left(w_{K}^{m}, 0 \right) \right\}}^{m} q_{w_{3-1}^{m}, w_{3}^{m}}^{2} + \sum_{k=1}^{K^{m}} \theta_{k}^{m} \right\}; (7)$$

минимизации времени завершения выполнения всех заданий:

$$F_{1}(\tilde{W}^{1}, ..., \tilde{W}^{m}, ..., \tilde{W}^{M}) = \min \max_{m=1, ..., M} \left\{ q_{0, w_{1}^{m}}^{1} + \frac{\sum_{(w_{9-1}^{m}, w_{9}^{m}) \in \{\tilde{W}^{m}/(0, w_{1}^{m})\}} q_{w_{9-1}^{m}, w_{9}^{m}}^{1} + \sum_{k=1}^{K^{m}} \theta_{k}^{m} \right\}, \quad (8)$$

$$F_{2}(\tilde{W}^{1}, ..., \tilde{W}^{m}, ..., \tilde{W}^{M}) = \min \max_{m=1, ..., M} \left\{ q_{w_{m}^{m}, 0}^{2} + \frac{\sum_{(w_{9-1}^{m}, w_{9}^{m}) \in \{\tilde{W}^{m}/(w_{k}^{m}, 0)\}} q_{w_{9-1}^{m}, w_{9}^{m}}^{2} + \sum_{k=1}^{K^{m}} \theta_{k}^{m} \right\}. \quad (9)$$

Обозначим $T[w_i^1|\tilde{W}]$, $T[w_i^2|\tilde{W}]$ — соответственно время первого и последнего посещения i-го пункта маршрута \tilde{W} . Тогда ограничения на выполнение сроков обслуживания пунктов могут быть представлены в виде

$$T[w_i^1|\tilde{W}] \geq h_i^1; \quad T[w_i^2|\tilde{W}] \leq h_i^2.$$

Назовем $\tilde{W}^s=\{0,\ i_1^s,\ i_2^s,\ ...,\ i_l^s,\ ...,\ i_p^s\}$ — некоторый частичный маршрут движения TC, для которого определены подмножества пунктов $\tilde{J}^{1,s}$ и $\hat{J}^{2,s};\ L^s=i_p^s$ — последний пункт s-го частичного плана в подпоследовательности \tilde{W}^s ; $\theta(\tilde{W}^s)$ — время завершения всех заданий в конечном пункте L^s в маршруте \tilde{W}^s , и этот маршрут должен быть продолжен дальше из этого пункта.

Далее рассматривается решение сформулированных задач с помощью последовательных алгоритмов оптимизации (модифицированными методами ветвей и границ и динамического программирования) на основе установленных оценок возможности выполнения всей системы ограничений и минимальных значений функции цели задачи, а также правил сравнения и определения предпочтений рассматриваемых частичных маршрутов и отсева бесперспективных продолжений.



2. ОЦЕНКИ ДОПУСТИМОСТИ И КРИТЕРИЕВ ОПТИМАЛЬНОСТИ ЧАСТИЧНЫХ ПЛАНОВ. ПРАВИЛА ОТСЕВА БЕСПЕРСПЕКТИВНЫХ ПРОДОЛЖЕНИЙ

Преобразуем матрицу $T = \|t_{ij}\|$ размерностью (n+1) в квадратную матрицу или матрицу $Q^{\beta} = \|q_{k,p}^{\beta}\|, k, p = 0, 1, ..., K, \beta = 1, 2$, размерностью (K+1), где значения $q_{k,p}^1$ вычисляются по формулам (1), а $q_{k,p}^2$, — по формулам (2). Как правило, (K+1) < (n+1), т. е. число всех пунктов сети превышает число пунктов приема грузов. Обозначим $Q^{1, s} = \|q_{k,p}^{1, s}\|, Q^{2, s} = \|q_{k,p}^{2, s}\|,$ — соответственно матрицы Q^1 и Q^2 на некотором s-м шаге преобразования. Размерность матриц $Q^{1, s}$ и $Q^{2, s}$ равна $K^{s, -} \le (K+1)$. При этом для s = 0 справедливо $Q^{1, 0} = Q^1$, $Q^{2, 0} = Q^2$.

Определим

$$g_k^{\beta,s}(0) = \min_{0 \le p \le K^s} q_{k,p}^{\beta,s}, \ \overline{q}_{k,p}^{\beta,s} = q_{k,p}^{\beta,s} - g_k^{\beta,s}(0),$$
$$\beta = 1, 2, \quad k = 1, ..., K^s, \tag{10}$$

$$\bar{g}_{p}^{\beta,s}(0) = \min_{0 \le k \le K^{s}} \bar{q}_{k,p}^{\beta,s}, \ d_{k,p}^{\beta,s} = \bar{q}_{k,p}^{\beta,s} - \bar{g}_{p}^{\beta,s}(0),$$
$$\beta = 1, 2, \quad k = 1, ..., K^{s}. \tag{11}$$

В полученных матрицах $G^{1, s} = \|\bar{g}_{k, p}^{1, s}\|$, $G^{2, s} = \|\bar{g}_{k, p}^{2, s}\|$ в каждой строке, а в матрицах $D^{1, s}$ или $D^{2, s}$ — в каждой строке и в каждом столбце содержится по крайней мере по одному элементу, значение которого равно нулю.

Определим сумму всех приводящих констант матриц $Q^{1, s}$ и $Q^{2, s}$:

$$\Delta^{\beta, s} = \sum_{k=1}^{K^s} g_k^{\beta, s}(0) + \sum_{p=1}^{K^s} \bar{g}_p^{\beta, s}(0), \quad \beta = 1, 2. (12)$$

Преобразованием 1 назовем преобразование матриц $Q^{1, s}$ или $Q^{2, s}$ в матрицы $G^{1, s}$, $D^{1, s}$ и $G^{2, s}$, $D^{2, s}$ по формулам (10), (11).

Отметим, что преобразование 1 выполняется на начальном этапе решения задачи 1 или 2 при s=0, т. е. для матриц $Q^{1,\;0}=\|q_{k\;p}^{1,\;0}\|$ или $Q^{2,\;0}=\|q_{k\;p}^{2,\;0}\|$.

Пусть на некотором шаге решения задачи в s-м частичном плане, когда осталось под-множество невыполненных заданий $\hat{J}^{2,s}$, которые отражены в матрицах $G^{1,s}$ или $G^{2,s}$, в качестве альтернативы

выбрано, что 9-е задание выполняется непосредственно перед заданием μ , где $J \in \hat{J}^{2,s}$ и $\mu \in \hat{J}^{2,s}$.

Преобразованием 2 назовем выполнение следующих вычислений:

- в матрицах $D^{1, s}$ или $D^{2, s}$ вычеркиваем 9-ю строку и μ -й столбец;
- элемент $d_{9\mu}^{1,\,s}$ или $d_{9\mu}^{2,\,s}$ матриц полагаем стремящимся к бесконечности;
- выполняем преобразование 1 вновь полученных матриц.

Пусть на некотором этапе решения задач 1 и 2 построен частичный маршрут выполнения заданий

$$\tilde{W}^{s} = \{w_0, w_1 ..., w_k, ..., w_{I^s}\}.$$
 (13)

Утверждение 1. Нижняя граница функции цели решения задач для s-го частичного плана решения задач 1 и 2 (выбора очередной пары выполняемых друг за другом заданий) определяется выражением

$$\xi\{T[\tilde{W}^{s}]\} = q_{0, w_{1}}^{\beta} + \sum_{(w_{9-1}, w_{9}) \in \{\tilde{W}^{s}/(w_{0}, w_{1})\}} q_{w_{9-1}, w_{9}}^{\beta} + \sum_{k=1}^{K} \bar{\tau}_{k} + \Delta^{\beta, s}, \quad \beta = 1, 2. \quad \bullet$$
(14)

Здесь и в дальнейшем $\tilde{W}^s/(w_0, w_1)$ — последовательность \tilde{W}^s , в которой исключена соответствующая пара пунктов (в данном случае (w_0, w_1)).

На начальном этапе решения задачи нижняя граница функции цели задачи рассчитывается по формуле

$$\xi\{T(\tilde{W})\} = \Delta^{\beta, s}, \quad \beta = 1, 2.$$
 (15)

Преобразуем матрицы $Q^{m,\;1,\;s}=\|q_{k,p}^{m,\;1,\;s}\|$ или $Q^{m,\;2,\;s}=\|q_{k,p}^{m,\;2,\;s}\|$ размерностью K^s каждая следующим образом

$$g_k^{\beta, s}(0) = \min_{1 \le m \le M} \min_{0 \le p \le K^s} q_{k, p}^{m, \beta, s},$$

$$\beta = 1, 2, \quad k = 1, ..., K^s;$$
(16)

$$\bar{q}_{k,p}^{m,\beta,s} = q_{k,p}^{m,\beta,s} - g_k^{\beta,s},$$

$$\beta = 1, 2, k, \quad p = 1, ..., K^s, \quad m = 1, ..., M; (17)$$

$$\bar{g}_p^{\beta,s}(0) = \min_{1 \le m \le M} \min_{0 \le p \le K^s} \bar{q}_{k,p}^{m,\beta,s},$$

$$d_{k,p}^{m,\beta,s} = \bar{q}_{k,p}^{m,\beta,s} - \bar{g}_{p}^{\beta,s}(0), \quad m = 1, ..., M.$$
 (18)



$$\overline{\Delta}^{\beta, s} = \sum_{k=1}^{K^{s}} g_{k}^{m, \beta, s}(0) + \sum_{p=1}^{K^{s}} \overline{g}_{p}^{m, \beta, s}(0),
m = 1, ..., M, \quad \beta = 1, 2.$$
(19)

Преобразованием 3 назовем преобразование матриц $Q^{m,\;1,\;s}=|q_{k,\;p}^{m,\;1,\;s}\|$ или $Q^{m,\;2,\;s}=|q_{k,\;p}^{m,\;2,\;s}\|$ в матрицы $G^{m,\;\beta,\;s}=\|g_{k,\;p}^{m,\;\beta,\;s}\|$ и $D^{m,\;\beta,\;s}=\|d_{k,\;p}^{m,\;\beta,\;s}\|,\;\beta=1,\;2,$ $m=1,\;...,\;M,$ в результате вычислений по формулам (16)—(18).

В результате преобразования 3 в каждой строке и каждом столбце, хотя в одной из матриц $D^{m, 1, s}$ или $D^{m, 2, s}$ содержится элемент, равный нулю. Сумма приводящих констант преобразования 3 определяется также по формуле (19).

Если на некотором этапе решения задач 3 и 4 построены частичные маршруты выполнения заданий на всех видах TC

$$\tilde{W}^{m,s} = \{w_0^{m,s}, w_1^{m,s}, ..., w_k^{m,s}, ..., w_{L^{m,s}}^{m,s}\},\$$

$$m = 1, ..., M.$$

Утверждение 2. Нижняя граница функции цели $F_1(\tilde{W}^1, ..., \tilde{W}^m, ..., \tilde{W}^M)$ решения задач на каждом s-м этапе решения задач 3 и 4 (выбора очередной пары выполняемых друг за другом заданий на одном из TC, m. e. $(w_{L^{m,s}}^{m,s}, k)$, $k \in \hat{K}^{s,-}$) определяется выражением

$$\xi\{F_{1}[\tilde{W}^{\beta,1,s},...,\tilde{W}^{\beta,M,s}]\} = \sum_{m=1}^{M} \left\{q_{0,w_{1}}^{m,\beta} + \sum_{(w_{9-1},w_{9}) \in \{\tilde{W}^{m,s}/(w_{0},w_{1})\}} q_{w_{9-1},w_{9}}^{m,\beta,s} + \sum_{w_{9} \in \tilde{K}^{m,s}} \theta_{w_{9}}\right\} + \bar{\Delta}^{\beta,s}; \quad \beta = 1, 2. \quad (20)$$

На начальном этапе решения этих задач нижняя граница функции цели задачи рассчитывается по формуле (15). Для критерия оптимальности (8), (9) нижняя граница функции цели определяется по формуле

Упорядочим все оставшееся подмножество невыполненных заданий $\hat{K}^{s,-}$ на некотором этапе решения задачи в последовательность по возрастанию граничных сроков завершения выполнения этих заданий

$$\tilde{U}\{\hat{K}^{s,-}\} = \{u_1, u_2, ..., u_{\gamma}, ..., u_{\hat{K}^s} | h_{u_{(\gamma-1)}}^2 \le h_{u_{\gamma}}^2, u_{\gamma} \in \hat{K}^{s,-}\}.$$
(22)

На начальном этапе решения задачи (s=0) эта последовательность имеет вид

$$\tilde{U}(\hat{J}^0 = \tilde{J}^2) = \{u_1, u_2, ..., u_k, ..., u_K | h_{u_{(k-1)}}^2 \le h_{u_k}^2, u_k \in \tilde{K}, k = 1, ..., K\}.$$
 (23)

Определим минимальные значения времен завершения выполнения заданий

$$\hat{g}_{p}^{\beta, s} = \min_{0 \le k \le \hat{K}^{s, -}} q_{k, p}^{\beta, s}, \quad \beta = 1, 2, \quad k \in \hat{K}^{s, -}.$$
 (24)

Пусть не некотором этапе решения задачи построен частичный маршрут выполнения заданий (13) — $k \in \tilde{W}^s$, $\lambda = w_{L^s}$ — индекс задания, выполняемого в этой последовательности последним. Время завершения этой последовательности заданий определяется по формуле

$$T[\tilde{W}^{s}] = q_{0, w_{1}}^{1} + \sum_{(w_{9-1}, w_{9}) \in \{\tilde{W}^{m, s}/(w_{0}, w_{1})\}} q_{w_{9-1}, w_{9}}^{1}.$$

Утверждение 3. Если для частичного плана \tilde{W}^s и подмножества подлежавших выполнению заданий $k \in \hat{J}^s$ справедливо хотя бы одно из системы неравенств

$$T[\tilde{W}^{s}] + q_{u_{L^{s}(p)}, u_{1}}^{1} > h_{u_{1}}^{2};$$

$$T[\tilde{W}^{s}] + q_{u_{L^{s}, u_{1}}} + \sum_{\rho=2}^{r} \hat{g}_{u_{\rho-1, p}}^{\beta, s} > h_{u_{\rho}}^{2},$$

$$r = 2, 3, ..., K^{s}, \qquad (25)$$

то подпоследовательность \tilde{W}^s не содержит допустимых решений. lacktriangle

$$\xi\{F_{2}[\tilde{W}^{\beta, m, s}, m = 1, ..., M]\} = \max \begin{cases} \min_{1 \leq m \leq M} \left[q_{0, w_{1}}^{m, \beta} + \sum_{(w_{9-1}, w_{9}) \in \{\tilde{W}^{m, s}/(w_{0}, w_{1})\}} q_{w_{9-1}, w_{9}}^{m, \beta, s} + \sum_{9=1}^{K^{m, s}} \theta_{9} \right] + \frac{1}{M} \overline{\Delta}^{\beta, s}; \\ \max_{1 \leq m \leq M} \left[q_{0, w_{1}}^{m, \beta} + \sum_{(w_{9-1}, w_{9}) \in \{\tilde{W}^{m, s}/(w_{0}, w_{1})\}} q_{w_{9-1}, w_{9}}^{m, \beta, s} \right]; \\ \beta = 1, 2. \end{cases}$$

$$(21)$$



Доказательство подобных утверждений можно найти в работах автора [3, 4, 24, 25].

Рассмотрим две подпоследовательности выполнения заданий \tilde{W}^s и \tilde{W}^δ , содержащие соответственно \tilde{K}^s и \tilde{K}^δ подмножества заданий. Пусть последним в каждой из этих подпоследовательностей стоит одно и то же задание w_9 , и времена выполнения этих подпоследовательностей соответственно равны $T[\tilde{W}^s]$ и $T[\tilde{W}^\delta]$.

Правило отсева неоптимальных продолжений 1. Если для двух подпоследовательностей выполнения заданий $\tilde{W}^s(p)$ и $\tilde{W}^\delta(\lambda)$ выполняются условия

$$\tilde{K}^{\delta} \subseteq \tilde{K}^{s} \text{ M } T[\tilde{W}^{s}] \leqslant T[\tilde{W}^{\delta}],$$
 (26)

то подпоследовательность \tilde{W}^s является предпочтительней подпоследовательности \tilde{W}^δ и последняя может быть исключена из рассмотрения как неперспективная.

3. АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ 2 МОДИФИЦИРОВАННЫМ МЕТОДОМ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Построим на основе матрицы $T = \|t_{ij}\|$ размерностью (n+1) квадратную матрицу $Q^1 = \|q_{k,\,p}^1\|$ или $Q^2 = \|q_{k,\,p}^2\|,\, k,\, p = 0,\, 1,\, ...,\, K$, размерностью (K+1), где значения $q_{k,\,p}^1$ вычисляются по формулам (1), а $q_{k,\,p}^2$ — по формулам (2).

Задача 1 решается с помощью алгоритмов решения задачи коммивояжера [22] с матрицами Q^1 или Q^2 , подробно описанных в работах [3—6], поэтому в настоящей статье на этом подробно останавливаться не будем.

В процессе решения задачи 2 модифицированными методами динамического программирования [26] и последовательного анализа вариантов производятся построение, анализ различных допустимых и конкурирующих маршрутов \tilde{W}^s движения ТС, начиная с начального пункта i=0, и развитие перспективных маршрутов до получения наиболее эффективного решения. В ходе построения производится отсев неперспективных частичных планов. На каждой итерации алгоритма (s=0,1,2,...,S), которая состоит из нескольких шагов, развивается перспективный план с наименьшим значением длины или нижней границы функции цели. Этот частичный план продолжается включением в мар-

шрут движения ТС одного из оставшихся альтернативных невыполненных заданий $k \in \hat{K}^{s,-}$, не включенных в подмножество \tilde{W}^s пунктов доставки грузов $j \in \hat{J}^{2s}$. Затем на основе правил (25) проверяется допустимость рассматриваемых альтернативных продолжений, осуществляются отсев неперспективных маршрутов на основе правил предпочтений (26) и переиндексация всех частичных планов. Первый допустимый план перевозок, значение критерия оптимальности которого меньше или равно значению суммарных затрат любого из конкурирующих и не продолженных до конца частичных планов, является решением задачи.

Приведенные в § 2 системы оценок и правила сравнения альтернативных частичных планов позволяют на ранних стадиях решения задачи отсеять строящиеся маршруты, не содержащие допустимых и оптимальных решений. Тот факт, что на каждой итерации развивается только строящийся маршрут с наилучшим значением функции цели, позволяет исключить анализ заведомо неоптимальных маршрутов.

Алгоритм 1. Метка каждого развивающегося частичного плана \tilde{W}^s содержит параметры: $\tilde{W}^s =$ $=\{w_0, w_1, ..., w_k, ..., w_{I^s}\}$ — подпоследовательность выполняемых заданий; $T[\tilde{W}^s]$ — время завершения выполнения этой подпоследовательности; w_{I^s} индекс пункта задания, выполняемого в этой последовательности последним \tilde{J}^{2s} и $\hat{J}^{s,-}$ — соответственно подмножество заданий, входящих и не включенных в последовательность \tilde{W}^s ; $\xi\{T[\tilde{W}^s]\}$ нижняя граница функции цели частичного плана $ilde{W}^s$; au_k^+ и au_k^- — соответственно время погрузки и разгрузки грузов k-го задания, $\tau_k^+ + \tau_k^- = \bar{\tau}_k$; $\psi\{\tilde{W}^s\}$ — параметр, определяющий, продолжен ли этот маршрут во все допустимые пункты — $\psi\{\tilde{W}^s\}=1$ или нет — $\psi\{\tilde{W}^s\}=0$, при $\psi\{\tilde{W}^s\}=2$ маршрут исключен из дальнейшего рассмотрения как недопустимый или как неоптимальный.

Пусть на некотором этапе решения задачи получено \tilde{Z}^r , r=1,...,R, допустимых маршрутов решения задачи, функции цели которых соответственно равны $F(\tilde{Z}^r)$, среди них $\overline{F}(\tilde{Z}^R) = \min_{1 \le r \le R} F(\tilde{Z}^r)$ — наилучшее среди всех полученных на данный момент решений задачи. В начале процесса решения



(s=0) формируем для частичного маршрута \tilde{W}^0 параметры начальной метки:

$$\tilde{W}^{0}(0) = \{0\};$$

$$\tilde{J}^{2,0} = \emptyset; \ \hat{J}^{0,-} = \tilde{K} = \{1, 2, ..., k, ..., K\};$$

$$T[\tilde{W}^{0}] = 0;$$

$$\xi\{T(\tilde{W}^0)\} = \sum_{k=2}^K \xi(q^1_{k-1,k}), \ \bar{F}(\tilde{Z}^R) = \infty; \ \ \psi\{\tilde{W}^0\} = 0.$$

Шаг 1. На каждой итерации алгоритма выполняем следующий объем вычислений. Среди всех частичных планов ($s=0,\ 1,\ 2,\ ...,\ S$) выбирается частичный план \tilde{U}^{η} , для которого

$$T[\tilde{W}^{\eta}] = \min_{0 \le s \le S} \{ T[\tilde{W}^s] | \psi(\tilde{U}^s) = 0 \}$$
 или $\xi \{ T[\tilde{W}^{\eta}] \} = \min_{0 \le s \le S} \{ \xi \{ T[\tilde{W}^s] \} | \psi(\tilde{U}^s) = 0 \}.$

Упорядочим все пункты приема и доставки грузов $\hat{J}^{\eta,-}$ в последовательности $\tilde{U}^{\eta}\{\tilde{J}^{s,-}\}$ в соответствии с выражениями (22) или (23). Вычислим минимальные значения времен завершения выполнения заданий $\hat{g}_{u_k}^{\beta,s}$ по формуле (24). Переходим к шагу 2.

Шаг 2. Рассматриваем частичный план \tilde{W}^{η} , последнее выполненное задание для которого — задание $\lambda = w_{L^{\eta}(p)}$. Для λ -го и каждого задания

 $k \in \hat{K}^{\eta,-}$ находим пары пунктов приема и доставки грузов $\pi_{\lambda} = (i_{\lambda}, j_{\lambda})$ и $\pi_{k} = (i_{k}, j_{k})$. Определим

$$\bar{T}\left[\tilde{W}^{\eta}\right] = \max\{T[\tilde{W}^{\eta}] + t(j_{\lambda}, i_{k}), h_{ki}^{1}\}.$$

Среди подмножества заданий $k \in \hat{K}^{\eta,-}$ находим такие, для которых выполняется система неравенств:

$$\begin{split} \overline{T}\left[\tilde{W}^{\eta}\right] + \; q_{\lambda,\,k}^{1,\,\eta} \leq h_{kj}^{2}, \; \overline{T}\left[\tilde{W}^{\eta}\right] + \; q_{\lambda,\,k}^{2,\,\eta} \leq h_{ki}^{2}, \\ k \in \; \tilde{U}^{\eta}\{\hat{K}^{\eta,\,-}\}; \end{split}$$

$$\bar{T}[\tilde{W}^{\eta}] + q_{\lambda, k}^{1, \eta} + \sum_{l \in \{\hat{K}^{\eta, -}/k\}} \xi(q_l^{1, \eta}) \le h_{lj}^2,$$

$$\bar{T}[\tilde{W}^{\eta}] + q_{\lambda,k}^{2,\eta} + \sum_{l \in \{\hat{K}^{\eta,-}/k\}} \xi(q_{li}^{2,\eta}) \le h_{li}^2,$$

$$l \in \tilde{U}^{\eta} \{\hat{K}^{\eta, -}\}, \quad l = 1, 2, ..., K^{\eta}, \quad l \neq k.$$

Пусть это будут подмножества заданий $k\in$ $\in \hat{K}_1^{\eta,-}\subseteq \hat{K}^{\eta,-}$ и пунктов $j\in \hat{J}_1^{\eta,-}\subseteq \hat{J}^{\eta,-}$.

- а. Если $\hat{J}_1^{\eta,-}=\varnothing$, т. е. таких пунктов не существует, то частичный план \tilde{U}^{η} не содержит допустимых решений и исключается из дальнейшего рассмотрения. В этом случае полагаем, что $\psi(\tilde{U}^{\eta})=2$, производим переиндексацию всех рассматриваемых частичных планов и переходим к шагу 1.
- б. Если $\hat{J}_1^{\eta,-}\neq\varnothing$, то для каждого из пунктов, $i_k\in(\hat{J}_1^{\eta,-}/\lambda),\ k=1,\ ...,\ K(\eta),$ где $K(\eta)$ число заданий этого подмножества, определяем:

$$\begin{split} T[\tilde{W}^{S+k}] &= T[\tilde{W}^{\eta} \cup i_{k}] = \max\{\bar{T} \, [\tilde{W}^{\eta}] \, + \, q_{\lambda, \, k}^{2} \, ; \, h_{ki}^{1}\}, \\ T[\tilde{W}^{S+k}] &= T[\tilde{W}^{\eta} \, \cup (i_{k}, j_{k})] = \\ &= \max(\max\{T[\tilde{W}^{\eta} \, \cup i_{k}] + \bar{\tau}_{k} \, + \, t(i_{k}, j_{k}); \, h_{ki}^{1}\}; \, h_{ki}^{2}), \\ k &= 1, \, ..., \, K(\eta). \end{split}$$

в. Полагаем $\psi(\tilde{U}^{\eta})=1$. Производим переиндексацию всех оставшихся для рассмотрения подмножеств и вычисляем нижние границы функции цели для них по формулам (14). Переходим к шагу 3.

Шаг 3. Продолжаем частичный план \tilde{W}^{η} во все пункты подмножества $i_{u_k} \in (\hat{J}_1^{\eta,-}/\lambda)$. При этом метка каждого из продолженных путей формируется следующим образом:

$$\begin{split} \tilde{W}^{S+k} &= \{0, \ u_1^{\eta}, \ u_2^{\eta}, \ ..., \ u^{\eta}, \ ..., \ u_{\lambda}^{\eta}, \ u_{k}\}; \\ u_{k}; \ T(\tilde{W}^{S+k}), \quad J^{2, \ s+k} &= (\tilde{J}^{2, \ S} \cup u_{k}); \\ \xi[T(\tilde{W}^{S+k})]; \ \psi(\tilde{W}^{S+k}) &= 0. \end{split}$$

Сравниваем каждый вновь полученный частичный план со всеми ранее построенными перспективными планами. Если для некоторой рассматриваемой пары маршрутов выполняются условия правила отсева неоптимальных продолжений 1, то соответствующий неперспективный частичный план исключается из дальнейшего рассмотрения, т. е. метка его $\psi(\tilde{W}^S)$ полагается равной 2.

Если для некоторого частичного плана выполняются условия $\tilde{K}^{S+k} = \tilde{K}$, то получено допустимое решение задачи. Полагаем $\psi(\tilde{W}^{S+k}) = 1$ и вычисляем время выполнения всего комплекса работ с учетом возврата TC в нулевой пункт

$$F(\tilde{W}^{S+k}) = T(\tilde{W}^{S+k}) + t(j_k, 0),$$

$$F(\tilde{W}^{S+k}) = T(\tilde{W}^{S+k}) + t_{j_{k}, 0}.$$



Если $F(\tilde{W}^{S+k}) < \bar{F}(\bar{Z}^R)$, то полагаем $\bar{F}(\bar{Z}^R) = F(\tilde{W}^{S+k})$, запоминаем этот маршрут $\tilde{Z}^R = \tilde{W}^{S+k}$ и переходим к шагу 4.

Шаг 4. Исключаем из дальнейшего рассмотрения все продолженные на этой итерации маршруты $\tilde{W}^{(S+r)}$, для которых справедливо неравенство $T[\tilde{W}^{(S+r)}] \geq F(\tilde{Z}^T)$ или $\xi\{T[\tilde{W}^{(S+r)}]\} \geq F(\tilde{Z}^T)$, положив для них $\psi(\tilde{W}^{(S+r)}) = 3$. Производим переиндексацию всех перспективных частичных планов, для которых $\psi(\tilde{U}^S) = 0$, пронумеровав их индексами s = 1, ..., S.

Если выполняются условия

$$\overline{F}(\tilde{Z}^R) \le \min_{1 \le s \le S} \{\xi[T(\tilde{W}^S)] | \psi(\tilde{W}^S) = 0, \quad (27)$$

то маршрут \bar{Z}^R со значением функции цели, равным $\tilde{F}(\bar{Z}^R)$, является решением задачи, и алгоритм завершает работу.

Если $S \neq 0$ и условия (27) не выполняются, то переходим к шагу 1. Если S=0 и $\bar{F}(\tilde{Z}^R)=\infty$, то не существует допустимых решений задачи, и алгоритм завершает работу.

Через некоторое число итераций алгоритма будет либо найдено оптимальное решение задачи, либо установлен факт, что исходная система ограничений задачи является несовместна.

В Приложении приведены иллюстративные примеры решения задач 1 и 2.

Отметим, что в случае большой размерности задачи могут применяться приближенные методы решения, когда на шаге 3 используются грубые правила отсева бесперспективных продолжений. Кроме того, определенный на первых итерациях эффективный маршрут может быть продолжен до конца (назовем это рекордом). При этом не следует обращать внимания на другие, более короткие конкурирующие маршруты, включающие меньшее число заданий и пунктов с лучшим значением критерия эффективности. Если отклонение критерия оптимальности от вычисленной нижней границы критерия в оптимальном решении значительно, то развивается следующий частичный план до тех пор, пока значение $T[\tilde{W}^{S+k}]$ остается лучше рекорда.

4. РЕШЕНИЕ ЗАДАЧ З И 4 ПОСЛЕДОВАТЕЛЬНЫМИ АЛГОРИТМАМИ ОПТИМИЗАЦИИ

Построим для каждого TC, m=1,...,M, на основании матриц $T^m=\|t_{ij}^m\|$ размерностью (n+1)

квадратные матрицы $Q^{m, \beta} = ||q_{k,p}^{m,\beta}||, \beta = 1, 2; k,$ p = 0, 1, ..., K, размерностью (K+1), где значения $q_{k,p}^m$ вычисляются по формулам (19)—(21).

Решение задачи 3 с критериями оптимальности (6)—(9) может быть получено методом ветвей и границ с помощью алгоритмов решения задачи M коммивояжеров с матрицами $Q^{m,\,\beta}$ описанными, например, в работе автора [23]. Оценка оптимального значения критерия оптимальности для каждого подмножества вариантов при этом вычисляется по формулам (20) и (21).

Алгоритм решения задачи 4 строится на основе модифицированных методов динамического программирования и последовательного анализа вариантов и состоит из двух этапов. При этом могут быть заданы ограничения на время работы ТС:

 $T[\tilde{W}^{m,s}] \leq H^m, \ m=1, ..., M$. На первом этапе решения для каждого из $TC, \ m=1, ..., M$, описанным выше алгоритмом 1 строятся допустимые и эффективные частичные маршруты с отсевом недопустимых и неэффективных планов. На втором этапе анализируются различные варианты объединения различных допустимых объединений частичных маршрутов всех или некоторого подмножества машин в общий план выполнения всего комплекса заданий и выбирается среди этих допустимых объединений план, обеспечивающий оптимальное значение критерия эффективности.

Рассмотрим некоторые свойства таких допустимых объединений, а также правила их сравнения и отсева бесперспективных продолжений.

Правило 1. Для двух частичных планов $\tilde{W}^{m,s}$ и $\tilde{W}^{\mu,s}$, содержащих подмножества выполненных заданий $\tilde{J}^{2,m,s}$ и $\tilde{J}^{2,\mu,s}$ соответственно, объединение в допустимое расписание $\Omega\{\tilde{W}^{m,s},\,\tilde{W}^{\mu,s}\}$ возможно, если $\tilde{J}^{2,m,s}\cap \tilde{J}^{2,\mu,s}=\emptyset$.

Правило 2. Если для некоторого объединения частичных маршрутов $\tilde{\Omega}\{\tilde{W}^{m_1,\,s},\,\tilde{W}^{m_2,\,s},\,...,\,\tilde{W}^{m_P,\,s}\},$ где $P \leq M$, выполнены условия

$$\begin{split} \tilde{J}^{2,\,m_{p},\,s} & \cap \ \tilde{J}^{2,\,m_{p},\,s} = \varnothing, \quad r \neq p,\,r, \quad p = 1,\,...,\,P; \\ & \bigcup_{p=1}^{P} \,\tilde{K}^{m_{p},\,s} = \,\tilde{K}\,, \end{split}$$

то получено допустимое расписание выполнения всего комплекса заданий на M машинах, значение критерия оптимальности которого вычисляется по формулам (6), (7) или (8), (9).



Правило отсева. Если для двух допустимых объединений частичных планов $\tilde{\Omega}^1\{\tilde{W}^{m_1,\,s},\,\tilde{W}^{m_2,\,s},\,...,\,\tilde{W}^{m_p,\,s}\}$ и $\tilde{\Omega}^2\{\tilde{W}^{m_1,\,\rho},\,\tilde{W}^{m_2,\,\rho},\,...,\,\tilde{W}^{m_p,\,\rho}\}$, для которых

$$\tilde{J}^{2}(\tilde{\Omega}^{1}) = \bigcup_{p=1}^{P} \tilde{J}^{2, m_{p}, s} \text{ M } \tilde{J}^{2}(\tilde{\Omega}^{2}) = \bigcup_{p=1}^{P} \tilde{J}^{2, m_{p}, \rho},$$

справедливы условия

$$\tilde{J}^2(\tilde{\Omega}^2) \subseteq \tilde{J}^2(\tilde{\Omega}^1), \ \tilde{M}(\tilde{\Omega}^1) \subseteq \tilde{M}(\tilde{\Omega}^2)$$

$$\text{M} \ F_a(\tilde{\Omega}^1) \le F_a(\tilde{\Omega}^2), \tag{28}$$

где $\alpha = 1$ или 2, $\tilde{M}(\tilde{\Omega})$ — подмножество маршрутов TC, которые включены в соответствующее объединение,

и одно из этих условий выполняется как строгое соотношение, то объединение $\tilde{\Omega}^1$ более предпочтительно, чем объединение $\tilde{\Omega}^2$, и последнее может быть отброшено как неперспективное.

При построении различных вариантов объединения последовательностей $\tilde{\Omega}^r$ выполнения заданий и включении в это объединение следующего ТС r-я метка строящегося объединения r=1,...,R содержит признаки:

 N_1^r — номер TC, включаемого в объединение $\tilde{\Omega}^r; \ \overline{N}_2^r = \{N_2^{r,\,1},\ N_2^{r,\,2},\ ...,\ N_2^{r,\,m},\ ...,\ N_2^{r,\,M}\}$ — номера меток каждого TC, включаемого в объединение $\tilde{\Omega}^r.$

Отметим, что если $N_2^{r,\,m}=0$, то подмножество и последовательность заданий, выполняемых на т-м TC, не включены в объединение $\tilde{\Omega}^r$; $\tilde{\textbf{\textit{K}}}^r = \tilde{\textbf{\textit{J}}}^2(\tilde{\Omega}^r)$ — подмножество выполненных заданий в объединении $\tilde{\Omega}'$; $\tilde{M}(\tilde{\Omega}')$ — подмножество ТС, маршруты которых включены в объединение $\tilde{\Omega}^r$; $\tilde{\Omega}^r \{\tilde{W}^{m_1,s},$ $\tilde{W}^{m_2,\,s},\,...,\,\tilde{W}^{m_P,\,s}\}$ — последовательности выполнения заданий на каждом TC в объединении $\tilde{\Omega}^r$; δ^r номера включаемых в объединение $\tilde{\Omega}^r$ TC; $F_{\alpha}(\tilde{\Omega}^r)$, $\xi[F_{\alpha}(\tilde{\Omega}')]$, где $\alpha=1$ или 2 — соответственно значение критерия оптимальности и нижней границы функции цели для объединения подпоследовательностей в $\tilde{\Omega}^r$; N_3^r — возможность и перспективность продолжения $\tilde{\Omega}^r$. При этом $N_3^r=0$, если данное объединение последовательностей может быть продолжено, и $N_3^r = 1$ — если объединение $\tilde{\Omega}'$ уже продолжено во все возможные направления, $N_3^r=2$, если $\tilde{\Omega}^r$ исключено из рассмотрения как неперспективное, $N_3^r=3$, если получено допустимое решение задачи, т. е. $\tilde{K}^r=\tilde{J}^2(\tilde{\Omega}^r)=\tilde{K}$.

Рассматриваются только допустимые и перспективные метки последовательностей выполнения заданий на каждом TC, для которых значение $\psi(\tilde{W}^S)=1$. Обозначим через F_{α}^{\min} , где $\alpha=1$ или 2, наилучшие значения соответствующих критериев оптимальности для допустимых расписаний на данном шаге алгоритма. Введем счетчик строящихся объединений R. В начале процесса решения полагаем r=0, $\tilde{\Omega}^0=\varnothing$, $N_1^r=0$, $\bar{J}^+(\tilde{\Omega}^0)=\varnothing$, $\delta^0=0$, $F_{\alpha}(\tilde{\Omega}^0)=0$, $N_3^0=0$, $F_{\alpha}^{\min}=\infty$, а также R=0.

Алгоритм 2. *Шаг 1.* Выбираем некоторую r-ю метку объединения подпоследовательностей, для которой $N_3^r = 0$. Если таких меток не существует, то переходим к шагу 4. В противном случае выполняем следующие вычисления.

1.1. Среди меток TC с индексом $m = (\delta^r + 1) \le M$, для которой $\tilde{\Omega}^r \{\tilde{W}^{m_1,r}, \tilde{W}^{m_2,r}, ..., \tilde{W}^{\delta^r,r}\}$, находим последовательности выполняемых заданий, удовлетворяющие условиям

$$\psi(\tilde{\boldsymbol{W}}^{m,s}) = 1, \quad \tilde{\boldsymbol{J}}^2(\tilde{\boldsymbol{\Omega}}^r) \cap \tilde{\boldsymbol{J}}^2(\tilde{\boldsymbol{W}}^{m,s}) = \varnothing.$$

Если таких меток не существует, то $\delta' := (\delta' + 1)$ и снова выполняем п. 1.1.

В противном случае переходим к шагу 2. Если m > M, то переходим к шагу 3.

Шаг 2. Полагаем

$$\tilde{\Omega}^{r+1} \{ \tilde{W}^{m_1, s}, \ \tilde{W}^{m_2, s}, ..., \ \tilde{W}^{\delta^r, s}, \ \tilde{W}^{(\delta^r+1), s} \},$$

$$\tilde{J}^2 (\tilde{\Omega}^{(r+1)}) = \tilde{J}^2 (\tilde{\Omega}^r) \cup \tilde{J}^2 (\tilde{W}^{m, s}), \quad \delta^r := (\delta^r + 1).$$

Вычисляем

$$F_1(\tilde{\Omega}^{(r+1)}) = F_1(\tilde{\Omega}^r) + T[\tilde{W}^{m,s}],$$

$$F_2(\tilde{\Omega}^{(r+1)}) = \max_{m \in \tilde{M}(\tilde{W}^{m,s})} T[\tilde{W}^{m,s}].$$

Если m < M, то $\delta^r := (\delta^r + 1)$ и вновь переходим к шагу 1.

В противном случае (m = M) выполняем следующий объем вычислений.

2.1. Если $\tilde{J}'(\tilde{\Omega}^{(r+1)}) = \tilde{K}$, то определяем $F_{\alpha}^{\min} = \min\{F_{\alpha}^{\min}; F_{\alpha}(\tilde{\Omega}^{(r+1)})\}, \alpha = 1$ или 2.

Далее $r := (R+1), N_3^r = 3$. Переходим к шагу 3.

2.2. Если $\tilde{J}^r(\tilde{\Omega}^{(r+1)})\subset \tilde{K}$, то $r\coloneqq (R+1)$ и переходим к шагу 1.



Шаг 3. Для объединения $\tilde{\Omega}^{(r+1)}$ и каждого из всех построенных ранее объединений $\tilde{\Omega}^{\rho}$, $\rho < (r+1)$, для которых $\psi(\tilde{\Omega}^{\rho}) = 0$ или $\psi(\tilde{\Omega}^{\rho}) = 1$, проверяем выполнение условий доминирования (28) в соответствии с правилом отсева. Если более перспективным является $\tilde{\Omega}^{(r+1)}$ ($\tilde{\Omega}^{(r+1)} > \tilde{\Omega}^{\rho}$), то полагаем $\psi(\tilde{\Omega}^{\rho}) = 2$, после чего переходим к п. 1.1 шага 1. Если $\tilde{\Omega}^{\rho} > \tilde{\Omega}^{(r+1)}$, то переходим к шагу 1.

Шаг 4. Если $F_{\alpha}^{\min} \neq \infty$, то получено оптимальное решение задачи, и алгоритм завершает работу. Если $F_{\alpha}^{\min} \neq \infty$, то допустимых решений задачи не существует, и с соответствующим сообщением алгоритм завершает работу.

Отметим, что в практических расчетах решения задачи по критериям минимизации времени выполнения расписания работ (для сокращения числа вариантов на первом этапе решения задачи) могут быть заданы вспомогательные дополнительные ограничения на времена работы $TC\ T[\tilde{W}^{m,s}] \leq H^m, m=1,...,M$. Это сократит также объем вычислений в алгоритме 2. Вычислительные эксперименты. Решались зада-

чи с числом выполняемых заданий 7—15 и числом

ТС не более четырех. Число пунктов приема и доставки грузов не превышало 30. Время решения всех задач на персональном компьютере с процессором Core i-5 Intel® Core™ i5-4200U четвертой модификации (до 2,6 ГГц) составляло не более 2,5 мин. Наибольшее число анализируемых вариантов и соответственно наибольший объем вычислений наблюдались в задачах без ограничений с небольшими отличиями во временах и стоимости перевозки грузов из одного пункта в другой. Наличие жесткой системы ограничений и большие разбросы во временах и стоимости доставки грузов позволяют уже на первых итерациях алгоритма исключить маршруты, не содержащие допустимых и оптимальных планов, и существенно сократить число анализируемых маршрутов. В задачах большой размерности при несущественных ограничениях на сроки доставки грузов и небольших отличиях в значениях элементов матриц $C^m = \|c_{ii}^m\|$ и $T^m = ||t_{ij}^m||, m = 1, ..., M$, могут быть применены приближенные методы решения задачи, использующие в алгоритмах метода ветвей и границ одностороннюю схему ветвления до получения «рекорда» или в методах динамического программирования приближенные правила отсева неперспективных продолжений. Кроме того, в методах динамического программирования используются оценки. Продолжаются частичные маршруты с наилучшими оценками и некоторые расписания не доводятся до конца, если получены решения со значением критерия эффективности, лучшим, чем эти оценки.

ПРИЛОЖЕНИЕ

Проиллюстрируем работу алгоритма решения задач 1 и 2 на примере. Матрица времен переезда между пунк-

Таблица 1 Матрица времен переезда между пунктами

№		Времена проезда между пунктами приема и доставки грузов												
	0	1	2	3	4	5	6	7	8	9	10	11	12	
0	8	10	20	18	30	15	27	35	22	8	17	40	25	
1	10	∞	14	18	25	36	10	41	24	32	16	20	40	
2	20	14	∞	36	15	24	23	40	27	16	19	32	21	
3	18	18	36	∞	20	17	26	8	19	22	25	13	26	
4	30	25	15	20	∞	18	26	25	11	17	28	30	24	
5	15	36	24	17	18	8	9	16	28	15	41	23	10	
6	27	10	23	26	26	9	∞	15	18	30	12	14	25	
7	35	41	40	8	25	16	15	8	36	15	28	12	18	
8	22	24	27	19	11	28	18	36	∞	25	16	29	32	
9	8	32	16	22	17	15	30	15	25	∞	14	28	31	
10	17	16	19	25	28	41	12	28	16	14	∞	9	17	
11	40	20	32	13	30	23	14	12	29	28	9	∞	25	
12	25	40	21	26	14	10	25	18	32	31	17	17	∞	

Таблица 2

Задания по доставке грузов

Параметры	Номера и содержание выполняемых заданий									
	1	2	3	4	5	6				
Пункты приема груза	1	2	4	7	8	12				
Пункты доставки груза	11	6	10	3	5	9				
Время погрузки и разгрузки $\overline{\tau}_k$	5	3	6	7	2	4				
Граничные сроки выполнения заданий	85	230	185	320	280	130				

Таблица 3
Матрица времен между выполнением двух стоящих в последовательности друг за другом заданий

№	ме	еобхо жду в	Нулевые значения строк					
	0	1	2	3	4	5	6	
0	8	10	20	30	35	22	25	10
1	65	∞	57	55	37	54	50	37
2	52	36	× ×	52	41	44	51	36
3	51	50	53	∞	62	50	51	50
4	37	33	51	35	∞	34	41	33
5	45	66	54	48	46	oc	40	40
6	43	67	51	52	50	60	∞	43
Нулевые	_	_	3	2	_	_	_	254
значения								

68



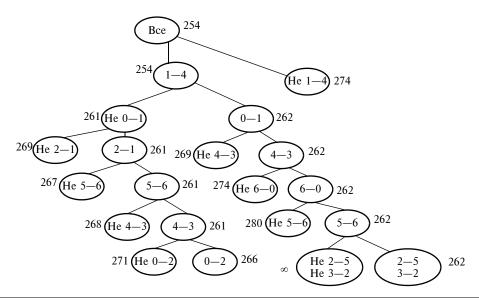


Рис. 1. Дерево решений задачи при отсутствии ограничений

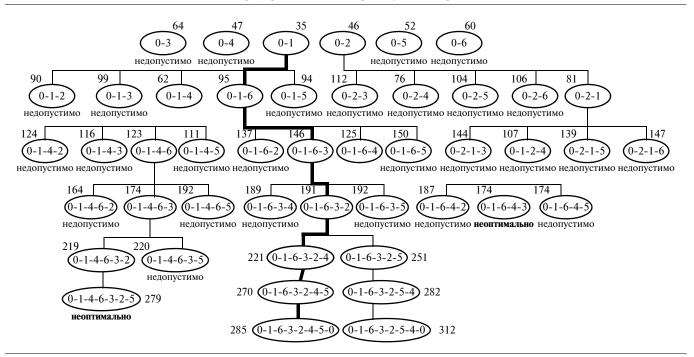


Рис. 2. Дерево решений задачи при наличии ограничений

Таблица 4

Приведенная матрица $D = Id_{kl}I$

					N/							
№	мех	Приведенная матрица времен проезда между парами пунктов выполняемых заданий										
	0	1	2	3	4	5	6					
0	∞	0	7	18	25	12	15					
1	28	∞	17	16	0	17	15					
2	16	0	∞	14	5	8	15					
3	1	0	0	∞	12	0	1					
4	0	0	15	0	∞	1	8					
5	5	26	11	6	6	∞	0					
6	0	24	5	7	7	17	∞					

тами дорожной сети представлена в табл. 1. Параметры шести заданий по доставке грузов приведены в табл. 2.

Например, расчет элемента матрицы $Q=|q_{k,\ p}|,$ $k,\ p=1,\ 2,\ ...,\ 6$, необходимого времени между выполнением двух рядом стоящих в последовательности заданий $q_{k,\ p}=q_{3,1}$ определяется выражением $q_{31}=\bar{\tau}_3+t_{4,10}+t_{10,1}=6+28+16=50;$ $q_{4,6}=\bar{\tau}_4+t_{7,3}+t_{3,12}=7+8+26=41.$ Рассчитанные элементы сведены в табл. 3. Приведенная матрица $D=|d_k|$ представлена в табл. 4.

Процесс решения примера в условиях отсутствия ограничений на сроки выполнения заданий методом ветвей и границ приведен на рис. 1.



Оптимальная последовательность выполнения заданий $\tilde{W}=\{0,1,4,3,2,5,6,0\}$, а последовательность пунктов сети при выполнении этого расписания $\tilde{V}=\{0,1,11,2,6,4,10,7,3,8,5,12,9,0\}$. Время выполнения расписания — 262.

Учет ограничений на сроки выполнения заданий. Пары пунктов приема и доставки грузов каждого задания определены в первой и второй строках, время, необходимое для загрузки и выгрузки груза *l*-го задания, — в третьей строке, а граничные сроки времен завершения выполнения заданий приведены в последней строке табл. 2. Необходимо найти допустимое расписание выполнения работ, минимизирующее время выполнения расписания.

Процесс решения примера методом динамического программирования с помощью алгоритма 1 приведен на рис. 2.

Внутри эллипсов показаны частные маршруты движения ТС. Цифры, стоящие рядом с вершиной дерева, показывают длину соответствующего частичного плана. На рис. 2 выделена ветвь дерева, соответствующая оптимальному решению.

Оптимальная последовательность выполнения заданий $\tilde{W}=\{0,1,6,3,2,4,5,0\}$, а последовательность пунктов сети при выполнении этого расписания $\tilde{V}=\{0,1,11,12,9,4,10,2,6,7,3,8,5,0\}$. Время выполнения расписания 285.

ЛИТЕРАТУРА

- 1. *Бенсон Д., Уайхед Дж.* Транспорт и доставка грузов: пер. с англ. М.: Транспорт, 1990. 279 с.
- 2. *Беленький А.С., Левнер Е.В.* Применение моделей и методов теории расписаний в задачах оптимального планирования на грузовом транспорте: Обзор // Автоматика и телемеханика. 1989. № 1. С. 3—77.
- 3. Зак $\emph{Ю.А.}$ Математические модели и алгоритмы построения эффективных маршрутов доставки грузов. М.: РУСАЙНС, 2015. 306 с.
- 4. Зак Ю.А. Об одной задаче построения допустимых и оптимальных маршрутов доставки грузов. Информационные технологии. 2015. № 5. С. 373—384.
- Domschke W. Logistik: Transport. Grundlagen lineare Transport- und Umladeprobleme.
 Auflage. Мъпсhen-Wien: R. Oldenburg Verlag, 2007. 234 s.
- Grünert T., Imich St. Optimierung in Transport, 1. Grundlagen & 2. Wege und Touren: Shaker Verlag, 2005.
- 7. *Jepsen M.*, *Petersen B.*, *Spoorendonk S.*, *Pisinger D.* Subset-row inequalities applied to the vehicle-routing problem with time windows // Operations Research. 2008. Vol. 56, N 2. P. 497—511.
- 8. *Bianchessi N., Righini G.* Heuristic algorithms for the vehicle routing problem with simultaneous pick-up and delivery // Computers & Operations Research. 2007. Vol. 34 (2). P. 578—594.
- 9. *Montano F.A.T.*, *Calvao R.D.* Vehicle Routing Problem with Simultaneous Pick-ap Up and Delivery Service // Operational Research of India. 2002. Vol. 39, N 1. P. 19—33.
- 10. Montano F.A.T., Calvao R.D. A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery

- service // Computers & Operations Research. -2006. Vol. 33. P. 595 619.
- 11. *Gan X., Yan W., Li Sh., Niu B.* Vehicle Routing Problem with Time Windows and Simultaneous Delivery and Pick-Up Service Based on MCPSO // Mathematical Problems in Engineering. 2012. Vol. 6. 11 p.
- 12. Shimizu Y., Sakaguchi T., Yoo J. A hybrid method for solving multi-depot VRP with simultaneous pickup and delivery incorporated with Weber basis saving heuristic // Journal of Advanced Mechanical Design, Systems, and Manufacturing. 2016. Vol. 10, N 1. P. 1—13.
- Ai J., Kachitvichyanukul V. A particle swarm optimization for the vehicle routing problem with simultaneous pickup and delivery // Computers & Operations Research. — 2009. — Vol. 36, N 5. — P. 1693—1702.
- 14. Gajpa Y., Abad P. An ant colony system (ACS) for vehicle routing problem with simultaneous delivery and pickup // Computers & Operations Research. — 2009. — Vol. 36, N 12. — P. 3215—3223.
- Wu C., Zhang N., Jiang J., Yang J., Liang Y. Improved bacterial foraging algorithms and their applications to job shop scheduling problems // ICANNGA, Part I. 2007. LNCS 4431. P. 562—569.
- 16. *Jeon G.*, *Leep H.R.*, *Shim J.Y.* A vehicle routing problem solved by using a hybrid genetic algorithm // Computers & Industrial Engineering. 2007. Vol. 53, N. 4. P. 680—692.
- Dash S., Gunluk O., Lodi A., Tramontani A. A time bucket formulation for the traveling salesman problem with time windows // INFORMS Journal on Computing. 2012. Vol. 24. P. 132—147.
- Baldacci R., Mingozzi A., Roberti R. Recent exact algorithms for solving the vehicle routing problem under capacity and time window constraints // European Journal of Operational Research. — 2012. — Vol. 218, iss. 1. — P. 1—6.
- Savelsbergh M. Local search in routing problems with time windows // Annals of Operations Research. 1985. Vol. 4, N 1. P. 285—305.
- Gehring H., Homberger J. Parallelization of a two-phase metaheuristic for routing problems with time windows // Journal of Heuristics. — 2002. — Vol. 8, N 3. — P. 251—276.
- Lin C.K.Y. A vehicle routing problem with pickup and delivery time windows, and coordination of transportable resources // Computers & Operations Research. 2011. Vol. 38, N 11. P. 1596—1609.
- Little J.D.C., Murty K.G., Sweeney D.W., and Karel C. An algorithm for the Traveling Salesman Problem // Operations Research. 1963. N 11. P. 972–989.
- 23. Зак Ю.А. Алгоритмы решения задач N коммивояжеров // Кибернетика. 1972. № 2. С. 96—106.
- 24. Зак Ю.А. Прикладные задачи теории расписаний и маршрутизации перевозок. М.: Либроком, 2012. 393 с.
- 25. Зак Ю.А. Распределение множества заданий и определение очередностей их выполнения на параллельных машинах методами динамического программирования. LAP LAMPERT Academic Publishing, 2013. 116 р.
- Беллман Р. Динамическое программирование. М.: Изд-во иностр. литературы, 1960.

Статья представлена к публикации членом редколлегии A.A. Лазаревым.

УДК 658.52.011.56

ЛОГИЧЕСКИЕ МОДЕЛИ ОБЪЕКТОВ ЖЕЛЕЗНОДОРОЖНОЙ СТАНЦИИ

А.И. Потехин

Рассмотрены проблемы создания систем управления движением поездов, способных предвидеть возможные конфликтные ситуации. С этой целью разработан формальный метод представления структуры железнодорожных объектов (станций, перегонов) в виде системы логических функций, на основе которых находится множество маршрутов. Определены логические отношения между маршрутами (совместимые, несовместимые, враждебные маршруты).

Ключевые слова: логическая модель, железнодорожная станция, стрелочный перевод, маршрут движения поездов, совместимые, враждебные маршруты.

ВВЕДЕНИЕ

Решение проблем повышения безопасности движения на железнодорожном транспорте связано с исследованием вопросов автоматизации диспетчерского управления движением поездов в железнодорожной сети. Основные функции диспетчерского управления состоят в оперативном контроле выполнения графика движения и формирования поездов и принятии решений по управлению движением поездов (при отклонении от графика).

Центральным звеном в оперативном управлении подвижным составом железнодорожной станции является станционный диспетчер (дежурный по станции). Диспетчер единолично распоряжается приемом, отправлением и пропуском поездов по главным и приемоотправочным путям железнодорожной станции. С помощью устройств сигнализации, централизации и блокировки он дистанционно управляет стрелками и сигналами, т. е. формирует поездные и маневровые маршруты; отдает приказы на движение поездов путем управления светофорами; при необходимости корректирует планы движения каждого поезда в режиме реального времени.

В периоды высокой интенсивности движения поездов (а также в нештатном режиме, в случае аварии) требуется высокая скорость принятия решения (к примеру, в Москве на Ярославском железнодорожном узле в «часы пик» прибывают и отправляются электропоезда каждые 2 мин).

В то время как расписание движения поездов тщательно планируется с помощью сложных ма-

тематических моделей, повседневная работа диспетчеров по выполнению план-графика движения поездов опирается только на свод правил, опыт и мастерство самих диспетчеров. Конечно, современные системы диспетчерского контроля и централизации (например, система «Сетунь» [1]) существенно облегчают труд диспетчеров. Однако при этом отсутствует какая-либо помощь диспетчеру в виде краткосрочного прогноза по процессу управления движением при наличии временных задержек поездов, неисправностей в инфраструктуре железнодорожной сети, конфликтных ситуаций. Отсутствуют отечественные средства оценки эффективности принимаемых решений. На практике отмечается, что 80 % всех случаев нарушений безопасности движения в перевозочном процессе, так или иначе, связаны с виной диспетчеров.

1. ПОСТАНОВКА ЗАДАЧИ

Коммуникационные, транспортные, производственные системы, отличающиеся недетерминированностью, асинхронностью, управлением по событиям, параллельностью в своих структурах, представляются как дискретно-событийные системы (discrete event systems — DES) [2—4].

В настоящее время во многих научных центрах мира ведутся широкие исследования по проблеме автоматизации управления движением поездов с целью разработать систему поддержки диспетчера, которая в состоянии предвидеть возможные конфликтные ситуации и находить решения, позволяющие избегать (исключать) эти ситуации.



Конфликтные ситуации могут возникать по разным причинам: опоздания прибытия или отправления поездов на станцию (временные помехи), неисправности элементов инфраструктуры станции (стрелок, путей, светофоров и др.), несогласованность группового движения поездов на станции и др.

Поэтому исследования, в основном, ориентированы на учет определенных типов помех. Так, в работах [5—8] исследуется проблема обеспечения безопасности группового движения поездов. Моделирование осуществляется на основе сетей Петри. Сети Петри по текущей разметке всегда точно определяют состояние железнодорожной сети, положение поездов, а существующие методы анализа сетей Петри позволяют определять (предвидеть) возможные конфликтные ситуации.

В работах [9—11] стратегии управления движением транспорта состоят в оперативной перепланировке маршрутов движения в режиме реального времени с целью оптимизации движения при наличии временных помех. Показано, что для этого необходимо уметь:

- определять (предвидеть) возможные конфликтные ситуации,
- осуществлять перепланировку маршрутов движения каждого поезда в режиме реального времени,
- оперативно осуществлять замену заданного маршрута альтернативным.

Для решения этих задач вне зависимости от принятой стратегии необходимо иметь:

- множество маршрутов объекта,
- множество альтернативных маршрутов заданному,
- множества совместимых, несовместимых, враждебных маршрутов заданному маршруту.

Предлагаемые решения эффективны только на локальном уровне при наличии небольших опозданий нескольких поездов. Поэтому во всех исследованиях основным приемом борьбы с размерностью служит декомпозиция больших задач на малые задачи, которые могут быть решены локальными диспетчерскими системами с дальнейшей их глобальной координацией.

Число маршрутов на крупных железнодорожных станциях может достигать 1000 и более. Однако большие временные затраты на выполнение процессов прогнозирования и перепланировки движения поездов обусловлены поиском альтернативных маршрутов, выбором оптимального маршрута по заданным критериям, проверкой выбранного маршрута на групповую совместимость с действующими маршрутами на железнодорожных объектах (станциях, перегонах).

Представляет научный и прикладной интерес исследование такой модели железнодорожной

станции, которая благодаря ее структурным особенностям сокращала бы перебор при решении указанных задач. Эти особенности состоят в дискретном характере объектов инфраструктуры станции, перегона. Например, путь (занят — свободен), стрелка (прямое — боковое положение, занята свободна) и др. Поэтому структуру железнодорожной станции, перегона можно рассматривать как релейно-контактную схему со многими входами и выходами или логическую схему из элементов ИЛИ, И, НЕ. Задача анализа логической схемы со многими выходами заключается в ее описании в виде системы логических функций [12, 13], которая позволяет находить полное множество маршрутов и служит исходной информацией для логических вычислений альтернативных маршрутов, групповой совместимости, враждебности маршрутов и других характеристик путем применения методов анализа и синтеза теории логических схем. Возможна аппаратная и программная реализации логических вычислений. По нашему мнению, этот путь исследования сократит временные затраты процессов прогнозирования и перепланировки движения поездов.

В настоящей статье рассматривается метод представления структуры железнодорожной станции в виде системы логических функций, на основе которой находится множество маршрутов станции и определяются логические отношения между маршрутами (совместимые, несовместимые, враждебные).

2. КРАТКОЕ ОПИСАНИЕ ОБЪЕКТА МОДЕЛИРОВАНИЯ

В железнодорожной сети России около 11 тыс. станций. Станцией называется «раздельный пункт» с путевым развитием и устройствами, позволяющими выполнять операции по приему, отправлению, скрещению и обгону поездов, а также по приему, погрузке, выгрузке и выдаче грузов и по обслуживанию пассажиров.

Основные характеристики станций без потери общности рассмотрим на примере схемы промежуточной железнодорожной станции. Основным компонентом станции является «именной» путь. Особенности именного пути: поезд при движении в прямом и обратном направлении проходит его полностью (без ответвлений), его границами могут быть стрелки и тупики, он свободен или занят только одним поездом (вне зависимости от фактической протяженности поезда). Именной путь это электрически изолированный рельсовый участок (блок-участок), состояние которого (свободен — занят) контролируется и управляется автоматически или диспетчером и отражается состоянием светофоров, расположенных на границах пути.



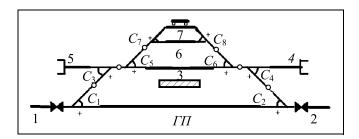


Рис. 1. Схема промежуточной железнодорожной станции

Каждый именной путь соединяется с помощью соединительного рельсового участка со стрелочными переводами. Стрелочный перевод служит для соединения, пересечения рельсовых путей и изменения направления движения поездов. По числу и расположению в топологическом плане пересекающихся путей стрелочные переводы могут быть одиночные, перекрестные и др. Основным элементом одиночного стрелочного перевода служит стрелка с переводным механизмом. Стрелки соединяются между собой и именными путями с помощью соединительных рельсовых участков. С помощью переводного механизма стрелка может находиться в одном из двух положений — прямом и боковом. На схемах станций знаком (+) обозначается прямое положение стрелок. Положение стрелки управляется и контролируется диспетчером. Стрелки, как и именные пути, являются электрически изолированными рельсовыми участками. Состояние стрелки (свободна — занята) управляется и контролируется.

На рис. 1 изображена схема станции, жирными линиями обозначены именные пути: Π — главный путь, пути I и 2 — вход-выходные пути станции, 3 — приемоотправочный путь, 4 и 5 — вытяжные пути (тупики), 6 — выставочный путь, 7 — погрузочно-разгрузочный путь, C_1 , ..., C_8 — стрелочные переводы (стрелки). Прямоугольником обозначена платформа посадки и высадки пассажиров (путь 3) и место погрузки-разгрузки (путь 7). Светофоры не показаны, чтобы не загромождать рисунок, их роль будет рассмотрена далее, управление ими либо автоматическое (блокировки), либо ручное (диспетчером).

3. ЛОГИЧЕСКАЯ МОДЕЛЬ ЖЕЛЕЗНОДОРОЖНОЙ СТАНЦИИ

3.1. Граф-схема железнодорожной станции

Несмотря на различные функциональные назначения станций, модель железнодорожной станции должна сохранить основные внешние свойства, как то: топологическую схему именных путей, схемы соединительных рельсовых участков и стрелочных переводов (стрелок). С этой целью исходную схему железнодорожной станции преобразуем в граф-схему специального вида. Это преобразование достаточно простое, покажем его на примере промежуточной станции.

Сформулируем правила преобразования.

- Именные пути станции (вне зависимости от их функционального назначения) на граф-схеме изобразим в виде кружков (рис. 2), которые в дальнейшем будем называть позициями, множество позиций обозначим как P, в нашем примере $P = \{p_1, p_2, , p_7\}$.
- Каждый именной путь станции связан со стрелками с помощью соединительных участков, которые изобразим в граф-схеме виде двух ребер, инцидентных позиции, соответствующей данному пути. Точки соединения этих ребер и позиций обозначим цифрами 1 и 2 (произвольно), они расположены вне кружков (см. рис. 2). Будем говорить, что эти точки соединения являются входами выходами позиции. Входы выходы *i*-й позиции обозначим как p_i^1 , p_i^2 .
- Стрелки станции в граф-схеме изображены в виде узлов, множество узлов обозначим как C, в нашем примере $C = \{c_1, c_2, ..., c_8\}$.
- Узлы соединяются между собой и позициями ребрами, которые соответствуют соединительным участкам станции. Знаком (+) обозначены ребра, которые указывают на прямое положение соответствующих стрелок. Ребра, соответствующие прямому и боковому положениям стрелки, соединены знаком \measuredangle . Множество ребер граф-схемы станции обозначим как U, оно состоит из неориентированных ребер типа $(p_i^1, c_i), (p_i^2, c_i)$ и (c_i, c_i) , где $p_i \in P$, $c_i, c_i \in C$.

Таким образом, граф-схема станции представляется в виде графа G, содержащего два типа вершин (множества P и C) и ребра (множество U): G = (P, C, U).

Преобразованная таким образом схема железнодорожной станции принимает вид, показанный на рис. 2. Число позиций в граф-схеме станции

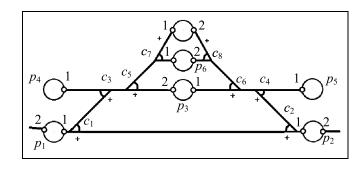


Рис. 2. Граф-схема промежуточной железнодорожной станции



можно изменить (увеличить, уменьшить). Например, при пересечении двух железнодорожных путей в одном уровне граф-схему станции необходимо дополнить позицией в месте пересечения. В то время как позицию, соответствующую главному пути станции, предназначенному для безостановочного следования поездов, можно опустить (с целью экономии числа позиций). В общем случае число и места расположения позиций определяются аналитиком.

Основным элементом организованного движения поездов на станции является маршрут. Простым станционным маршрутом будем называть простой путь в графе G в виде последовательности вершин и ребер:

$$m (p_i, p_j) = (p_i(p_i, c_{r, 1}), c_{r, 1}, (c_{r, 1}, c_{r, 2}), c_{r, 2}, ..., c_{r, k}, (c_{r, k}, p_j) p_j),$$

где позиции p_i и p_j соответствуют началу и концу маршрута, $(c_{r,\,\,1},\,c_{r,\,\,2},\,...,\,c_{r,\,\,k})$ —последовательность стрелок (узлов графа G), соединенных между собой соединительными участками станции (ребрами графа G): $(p_i,\,c_{r,\,\,1}),\,(c_{r,\,\,1},\,c_{r,\,\,2}),\,...,\,(c_{r,\,\,k},\,p_j)$.

Свойство простого маршрута: в нормальных ситуациях предполагается, что поезд проходит маршрут от начала до конца без остановки. Составным маршрутом будем называть маршрут, состоящий из нескольких простых маршрутов.

Не всякий путь в графе G удовлетворяет описанному свойству маршрута вследствие особенности движения поезда через стрелку. А именно, поезд физически не может перейти без остановки из прямого положения (направления) стрелки на боковое направление, а также — с бокового на прямое.

Формирование нового маршрута и организация движения поезда осуществляется диспетчером станции. При этом возможно возникновение проблем вследствие опозданий поезда, неисправностей в инфраструктуре железнодорожной сети, конфликтных ситуаций при групповом движении поездов на станции и др. Требуются большие временные затраты для прогнозирования и перепланировки движения поезда, поиска альтернативных маршрутов, выбора оптимального маршрута по заданным критериям, проверки выбранного маршрута на групповую совместимость с действующими маршрутами на станции.

Как будет показано далее, представление структуры железнодорожной станции в виде системы логических функций позволяет сократить перебор при нахождении множества маршрутов станции. Будут разработаны:

- логическая модель стрелочного перевода;
- логическая модель маршрута;

- метод нахождения маршрутов железнодорожной станции;
- метод анализа маршрутов на совместимость, враждебность.

3.2. Логическая модель стрелочного перевода

Вначале рассмотрим стрелочный перевод (стрелку), с помощью которого осуществляется управление направлением движения поездов. На рис. 3 приведена схема обыкновенной одиночной стрелки C. Вершина a — основной вход/выход стрелки, вершина b — прямой вход/выход, вершина d — боковой вход/выход, вершина e — центр стрелки. Положение стрелки описывается значением логической переменой s: s соответствует прямому положению (направлению) стрелки (на рис. 1 и 2 оно обозначено знаком (+)), положение \overline{S} — боковому направлению стрелки.

Схему стрелки можно рассматривать как логическую схему: в узле e реализуется логическая операция «проводное ИЛИ», ребра (b, e) и (d, e) взвешены логической переменной s, ребро (e, a) константой 1. Тогда, например, ориентированный путь [b, a] (b -начало, a -конец пути) в схеме стрелки будет проходным (по аналогии — электрическая проводимость релейной контактной цепи), если s = 1. С другой стороны, путь [b, a] будем считать разрешенным, если разрешено движение поезда на вход b, для чего введем логическую переменную x(b). Таким образом, будем считать, что путь [b, a] разрешенный и проходной для движения поезда, если x(b) = 1 и s = 1, что соответствует единичному значению конъюнкции sx(b). Проходимость разрешенного пути [b, a] представим в виде функции проходимости f(b, a):

$$f(b, a) = sx(b), \tag{1}$$

аналогично находим функции проходимости других путей:

$$f(d, a) = \overline{s}x(d),$$

$$f(a, b) = sx(a),$$

$$f(a, d) = \overline{s}x(a),$$

$$f(b, d) = x(b)s\bar{s} = 0, \quad f(d, b) = s\bar{s}x(d) = 0.$$
 (2)

Видно, что проходимость путей [b, d] и [d, b] невозможна.

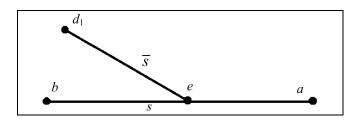


Рис. 3. Схема стрелки С



Обозначим через f(a) функцию проходимости на выходе a стрелки. Функция f(a) очевидным образом получается дизъюнкцией функций f(b, a) и f(d, a):

$$f(a) = sx(b) \vee \bar{s} x(d),$$

аналогично:

$$f(b) = f(a, b) = sx(a),$$

 $f(d) = f(a, d) = \bar{s}x(a).$ (3)

3.3. Логическая модель последовательности стрелок

Рассмотрим задачу получения результирующей функции проходимости последовательного соединения двух стрелок. Для ее решения применим известный принцип получения выходной функции заданной логической схемы: «от выхода схемы — к входам». Для этого рассмотрим процесс получения результирующей функции проходимости на выходе b_1 схемы «двойной стрелки» (рис. 4).

Шаг 1. Двигаясь от заданного выхода схемы (b_1) , получаем, что функция проходимости пути $[a_1, b_1]$ стрелки c_1 (согласно выражению (2)) $f(b_1) = s_1 x(a_1)$.

Шаг 2. Вход a_1 стрелки c_1 соединен с выходом a_2 стрелки c_2 (ребро (a_1, a_2) соответствует соединительному участку). Функция проходимости на выходе a_2 (согласно выражению (3)) $f(a_2) = s_2 x(b_2) \vee \overline{s}_2 x(d_2)$.

Шаг 3. Получение результирующей функции проходимости на выходе b_1 : отождествим переменные $x(a_1)$ и $f(a_2)$, получим $f(b_1) = s_1 f(a_2)$; подставим правую часть функции $f(a_2)$ в функцию $f(b_1)$, получим: $f(b_1) = s_1 [s_2 x(b_2) \vee \bar{s}_2 x(d_2)]$. Функцию $f(b_1)$ представим в виде дизъюнктивной нормальной формы (ДНФ):

$$f(b_1) = s_1 s_2 x(b_2) \vee s_1 \overline{s}_2 x(d_2). \tag{4}$$

В п. 3.5 будет рассмотрено получение результирующей функции проходимости для других вариантов последовательного соединения стрелок. Анализ ДНФ функции проводимости $f(b_1)$ представляет интерес, потому что на примере этой простой формулы можно предварительно сформулировать основные свойства функции проводимости любой сложности.

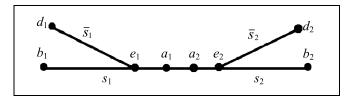


Рис. 4. Схема двойной стрелки

Сформулируем свойства функции проводимости $f(b_1)$.

- Число конъюнкций в ДНФ функции $f(b_1)$ равно числу путей, оканчивающихся в вершине b_1 (в нашем примере два пути).
- Каждая конъюнкция функции $f(b_1)$ содержит информации о начале и конце соответствующего пути. Так, переменная $x(b_2)$, помимо сигнала разрешения, указывает на путь с началом в вершине b_2 , переменная $x(d_2)$ путь с началом в вершине d_2 .
- Каждая конъюнкция функции $f(b_1)$ содержит информации о числе стрелок и их положении в каждом пути.
- Конъюнкции функции $f(b_1)$ допускают перестановку переменных (в силу коммутативного закона булевой алгебры $(s_1s_2 = s_2s_1)$). В то время как всякий путь имеет фиксированную последовательность стрелок. Таким образом, ДНФ функция $f(b_1)$ явно не содержит информации о последовательности стрелок в путях. Для того чтобы конъюнкции функции $f(b_1)$ содержали информацию о последовательности стрелок в путях, поступим следующим образом. Первой стрелке на шаге 1 рассмотренного процесса получения функции $f(b_1)$ присвоим порядковый номер 1. В нашем случае первой стрелкой является стрелка c_1 . Поэтому логической переменной s_1 присвоим порядковый номер 1 в виде верхнего индекса: s_1^1 . Последовательно со стрелкой c_1 соединена стрелка c_2 . Логической переменной s_2 присвоим порядковый номер 2 в виде верхнего индекса: s_2^2 . Тогда ДНФ функции $f(b_1)$ запишется как

$$f(b_1) = s_1^1 s_2^2 x(b_2) \vee s_1^1 \overline{s_2^2} x(d_2).$$

В этом случае каждая конъюнкция функции $f(b_1)$ содержит информацию о последовательности и одновременно о положении стрелок соответствующего пути.

Так, первой конъюнкции соответствует путь (b_2, s_2, s_1, b_1) , второй конъюнкции соответствует путь (d_2, \bar{s}_2, s_1, b) .

3.4. Логическая модель маршрута

Маршрут формируется диспетчером станции, он задает входную, выходную позиции маршрута и последовательность соединяющих их стрелок:

$$(p_i, p_i) = (p_i, c_{r, 1}, c_{r, 2}, ..., c_{r, k}, p_i),$$



где p_i — начало, p_j — конец маршрута. Здесь маршрут рассматривается как упорядоченная совокупность позиций и стрелок.

Далее, если все стрелки свободны (т. е. не заняты в других маршрутах), диспетчер устанавливает их в положение, соответствующее данному маршруту, которое задается набором, например: (1, 0, ..., 1). В наборе «1» означает прямое направление, «0» — боковое направление стрелки c. Положение каждой стрелки описывается значением логической переменой s. Единичное значение конъюнкции $(s_{r, 1}, \bar{s}_{r, 2}...s_{r, k})$ означает, что все стрелки последовательности $(c_{r,\ 1},\ c_{r,\ 2},\ ...,\ c_{r,\ k})$ установлены в положения, соответствующие набору (1, 0, ..., 1), и маршрут считается сформированным (собранным). Важное ограничение состоит в том, что маршрут может быть расформирован диспетчером либо до начала движения по нему поезда, либо после окончания движения, т. е. запрещается изменять маршрут во время движения по нему поезла.

Затем путем управления входными, выходными светофорами диспетчер дает разрешение на движение поезда по маршруту. Обозначим как $x(p_i)$ логическую переменную, соответствующую сигналу светофора, разрешающего выезд поезда с позиции p_i (начало маршрута), аналогично, $y(p_j)$ — разрешение на въезд поезда на позицию p_j (окончание маршрута). Тогда собранный маршрут $m(p_i, p_j) = (p_i, s_{r,1}, \bar{s}_{r,2}...s_{r,k}, p_i)$ считается разрешенным.

Функция проходимости разрешенного маршрута имеет вид:

$$f(p_i) = y(p_i)s_{r, k}...\bar{s}_{r, 2}, s_{r, 1}, x(p_i).$$

Функция $f(p_j)$, в отличие от функции (1), зависит от переменной $y(p_j)$ — разрешение на въезд поезда на позицию p_i (окончание маршрута).

3.5. Нахождение простых маршрутов железнодорожной станции

Станцию (см. рис. 2) представляем как логическую схему со многими входами и выходами, роль которых играют позиции. Пусть P — множество позиций станции, k — число стрелок. Для каждой позиции $p_i \in P$, считая ее выходной позицией, а остальные позиции — входными, специальной процедурой (от «выхода схемы — к входам») находим логическую функцию проходимости $f(p_i)$. Она принимает единичное значение на множестве разрешенных маршрутов, оканчивающихся в позиции p_i . Каждый разрешенный маршрут представляется конъюнкцией. Таким образом, функция

проходимости $f(p_i)$ равна дизьюнкции соответствующих конъюнкций.

Важным свойством функции проходимости $f(p_i)$, заключается в том, что между каждой конъюнкцией и соответствующим ей маршрутом имеет место взаимно однозначное соответствие (см. п. 3.3). По функции $f(p_i)$ получаем множество маршрутов станции $M(p_i)$, оканчивающихся в позиции p_i . Множество $\mathsf{U}M(p_i)$ по всем $p_i \in P$ определяет (задает) множество маршрутов станции.

Рассмотрим процесс построения маршрутов на примере построения маршрутов, оканчивающихся в позиции p_1 (см. рис. 2).

Шаг 1. Обозначим как $f(p_1^1)$ логическую функцию проходимости, принимающей единичное значение на множестве разрешенных маршрутах, оканчивающихся позицией p_1 входом 1:

$$f(p_1^1) = y(p_1^1)f(a_1),$$
 (5)

где $y(p_1^1)$ — разрешение въезда на позицию p_1 с входа $1, f(a_1)$ — функция проходимости на выходе a_1 стрелки c_1 .

Вход p_1^1 соединен с выходом a_1 стрелки c_1 , поэтому считаем, что стрелка c_1 — стрелка первого порядка.

На выходе a_1 стрелки c_1 (согласно выражению (3)) реализуется функция проходимости $f(a_1)$, и с учетом первого порядка стрелки c_1 имеем:

$$f(a_1) = s_1^1 x(b_1) \vee \overline{s}_1^1 x(d_1).$$

Вход b_1 стрелки c_1 соединен с выходом b_2 стрелки c_2 , вход d_1 — с выходом b_3 стрелки c_3 (см. рис. 2). Таким образом, стрелки c_2 и c_3 соединены со стрелкой c_1 , следовательно, эти стрелки второго порядка.

Согласно формуле (4)

$$f(a_1) = s_1^1 f(b_2) \vee \bar{s}_1^1 f(b_3),$$
 (6)

где $f(b_2)$ — функция проходимости на выходе b_2 стрелки c_2 , $f(b_3)$ — функция проходимости на выходе b_3 стрелки c_3 .

Шаг 2. В соответствии с рис. 2 строим функцию $f(b_2)$, которая согласно формулам (3), с учетом порядка стрелки c_2 , имеет вид:

$$f(b_2) = s_2^2 x(p_2^1), (7)$$

где $x(p_2^1)$ — логическая переменная, соответствующая сигналу светофора, разрешающего выезд по-



езда с входной позиции p_2 через выход 1, т. е. p_2^1 — начало маршрута.

Шаг 3. Функция $f(b_3)$, согласно формулам (3), строится аналогично: $f(b_3) = s_3^2 x(a_3)$, $a_3 = a_5$, в результате

$$f(b_2) = s_3^2 f(a_5), (8)$$

где $f(b_3)$ — функция проходимости на выходе a_5 стрелки c_5 , соединенной со стрелкой c_3 . Следовательно, стрелка c_5 — стрелка третьего порядка.

Шаг 4. В соответствии с рис. 2 строим функцию $f(a_5)$. Согласно формуле (4)

$$f(a_5) = s_5^3 x(p_3^2) \vee \bar{s}_5^3 f(a_7), \tag{9}$$

где $x(p_3^2)$ — логическая переменная, соответствующая сигналу светофора, разрешающего выезд поезда с позиции p_3 через выход 2, т. е. p_3^2 — начало маршрута.

Далее $f(a_7)$ — функция проходимости на выходе a_7 стрелки c_7 , т. е. стрелки четвертого порядка.

Шаг 5. В соответствии с рис. 2 строим функцию $f(a_7)$. Согласно формуле (4)

$$f(a_7) = s_7^4 x(p_7^1) \vee \bar{s}_7^4 x(p_6^1), \tag{10}$$

где $x(p_7^1)$, $x(p_6^1)$ — логические переменные, разрешающие выезд поезда с позиций p_7 , p_6 соответственно через выход 1, т. е. p_7^1 , p_6^1 — начала маршрутов.

Таким образом, в результате построены функции проводимости стрелок, определены порядки стрелок и начальные позиции всех маршрутов, оканчивающихся в позиции p_1 (см. рис. 2).

Следующий этап процесса построения маршрутов состоит в последовательности подстановок.

Шаг 6. Функция $f(a_5)$ строится путем подстановки правой части формулы (10) в формулу (9):

$$f(a_5) = s_5^3 x(p_3^2) \vee \bar{s}_5^3 [s_7^4 x(p_7^1) \vee \bar{s}_7^4 x(p_6^1)].$$
 (11)

Далее подставляем правую часть формулы (11) в формулу (8), затем правые части формул (7) и (8) — в формулу (6), правую часть формулы (6) — в формулу(5).

В результате подстановок получим скобочную форму функции $f(p_1^1)$:

$$f(p_1^1) = y(p_1^1)[s_1^1 s_2^2 x(p_2^1) \vee \bar{s}_1^1 s_3^2 [s_5^3 x(p_3^2) \vee \bar{s}_5^3 [s_7^4 x(p_7^1) \vee \bar{s}_7^4 x(p_6^1)]]].$$

Дизъюнктивная нормальная форма функции $f(p_1^1)$ имеет вид:

$$f(p_1^1) = y(p_1^1) s_1^1 s_2^2 x(p_2^1) \vee y(p_1^1) \bar{s}_1^1 s_3^2 \bar{s}_5^2 s_5^3 x(p_3^2) \vee y(p_1^1) \bar{s}_1^1 s_3^2 \bar{s}_5^3 \bar{s}_5^4 x(p_1^1) \vee y(p_1^1) \bar{s}_1^1 s_3^2 \bar{s}_5^2 \bar{s}_7^4 x(p_6^1).$$
(12)

Таким образом, функция $f(p_1^1)$ содержит четыре конъюнкции. Следовательно, множество маршрутов, оканчивающихся в позиции p_1^1 , содержит четыре простых маршрута.

Маршрут с началом в позиции p_2^1 и окончанием в позиции p_1^1 обозначим как $m(p_2^1, p_1^1)$. Первой конъюнкции функции $f(p_1^1)$ соответствует (собранный) маршрут

$$m(p_2^1, p_1^1) = (p_2^1, s_2, s_1, p_1^1).$$

Аналогично: $m(p_3^2, p_1^1) = (p_3^2, s_5, s_3, \bar{s}_1, p_1^1),$ $m(p_7^1, p_1^1) = (p_7^1, s_7, \bar{s}_5, s_3, \bar{s}_1, p_1^1),$ $m(p_6^1, p_1^1) = (p_6^1, \bar{s}_7, \bar{s}_5, s_3, \bar{s}_1, p_1^1).$

Аналогично определяются маршруты, оканчивающиеся в других позициях, а именно, число маршрутов, оканчивающихся в позиции p_2^1 , равно 4, в позиции p_3^1-2 , в позиции p_3^2-2 , в позиции p_4^1-3 , в позиции p_5^1-3 , в позиции p_6^1-2 , в позиции p_6^2-2 , p_7^1-2 , в позиции p_7^2-2 .

Функция $f(p_i) = f(p_i^1) \lor f(p_i^2)$, тогда в позиции p_1 оканчиваются 4 маршрута, в позиции $p_2 - 4$, в позиции $p_3 - 4$, в позиции $p_4 - 3$, в позиции $p_5 - 3$, в позиции $p_6 - 4$ и в позиции $p_7 - 4$ маршрута.

Итого имеем 26 простых маршрутов.

Нетрудно показать, что конъюнкции, входящие в ДНФ функции $f(p_i)$ для всех i, попарно ортогональны.

Можно показать, что:

- число конъюнкций, входящих в ДНФ функции $f(p_i)$, не более k_i+1 , где k_i суммарное число стрелок, входящих в маршруты, оканчивающихся в позиции $p_i,\ k_i \leq k;$
- число шагов процесса получения ДНФ функции $f(p_i)$, включая подстановки, не более $3k_i + 1$.



3.6. Нахождение составных маршрутов железнодорожной станции

Составной маршрут состоит из последовательности двух и более простых маршрутов. Основное свойство составного маршрута (как и простого) заключается в возможности его прохождения без остановки.

Из формулы (12) следует, что промежуточными позициями могут быть позиции $p_2,\,p_3,\,p_6,\,p_7.$

Обозначим как $h(p_1^1, p_3)$ функцию проходимости, которая принимает единичное значение на множестве составных маршрутов, оканчивающихся выходной позицией p_1^1 , содержащей позицию p_3 в качестве промежуточной.

Функция $h(p_1^1, p_3) = y(p_1^1)\bar{s}_1^1 s_3^2 s_5^3 x(p_3^2) f(p_3^1)$, где $f(p_3^1)$ — функция проходимости, которая принимает единичное значение на множестве простых маршрутов, оканчивающихся выходной позицией p_3^1 .

В соответствии с рис. 2 и с помощью изложенного процесса (см. п. 3.5) построим ДНФ функции $f(p_3^1)$: $f(p_3^1) = y(p_3^1)s_6^1s_4^2x(p_4^1) \vee y(p_3^1)s_6^1\bar{s}_4^2\bar{s}_2^3x(p_2^1)$. Подставим в функцию $h(p_1^1,p_3)$ правую часть функции $f(p_3^1)$, получим ДНФ функции $h(p_1^1,p_3)$:

$$h(p_1^1, p_3) = y(p_1^1)\bar{s}_1^1 s_3^2 s_5^3 x(p_3^2) y(p_3^1) s_6^1 s_4^2 x(p_4^1) \vee y(p_1^1)\bar{s}_1^1 s_3^2 s_5^3 x(p_3^2) y(p_3^1) s_6^1 \bar{s}_4^2 \bar{s}_2^3 x(p_2^1)].$$

Она содержит две конъюнкции. Следовательно, имеем два составных маршрута с началом в позициях p_4^1 , p_2^1 , концом в позиции p_1^1 и промежуточной позиции p_3 :

$$m(p_4^1, p_3, p_1^1) = (p_4^1, s_4, s_6, p_3, s_5, s_3, \bar{s}_1, p_1^1),$$

$$m(p_2^1, p_3, p_1^1) = (p_2^1, \bar{s}_2, \bar{s}_4, s_6, p_3, s_5, s_3, \bar{s}_1, p_1^1).$$

Аналогично вычисляются функции $h(p_1^1, p_2)$, $h(p_1^1, p_6)$, $h(p_1^1, p_7)$.

Функция $h(p_1^1, p_2) = 0.$

Составные маршруты с промежуточными позициями p_6, p_7 :

$$m(p_4^1, p_7, p_1^1) = (p_4^1, s_4, \bar{s}_6, s_8, p_7, s_5, s_3, \bar{s}_1, p_1^1),$$

$$m(p_2^1, p_7, p_1^1) = (p_2^1, \bar{s}_2, \bar{s}_4, \bar{s}_6, s_8, p_7, s_5, s_3, \bar{s}_1, p_1^1),$$

$$m(p_4^1, p_6, p_1^1) = (p_4^1, s_4, \bar{s}_6, \bar{s}_8^1, p_6, s_5, s_3, \bar{s}_1, p_1^1),$$

$$m(p_2^1, p_6, p_1^1) = (p_2^1, \bar{s}_2, \bar{s}_4, \bar{s}_6, \bar{s}_8^1, p_6, s_5, s_3, \bar{s}_1, p_1^1).$$

Аналогично определяются составные маршруты, оканчивающихся в других позициях. В результате имеем число составных маршрутов, оканчивающихся в позиции p_1-6 , в позиции p_2-6 , в позиции p_3-0 , в позиции p_4-6 , в позиции p_5-6 , в позиции p_6-0 , в позиции p_7-0 . Итого имеем 24 составных маршрутов.

3.7. Нахождение альтернативных маршрутов

Рассмотрим идею получения альтернативных маршрутов заданному маршруту $m(p_2^1, p_1^1) = (p_2^1, s_2, s_1, p_1^1)$, которому соответствует конъюнкция $y(p_1^1)s_1^1s_2^2x(p_2^1)$ функции $f(p_1^1)$ (формула (12)).

Альтернативные маршруты этому маршруту могут быть либо простыми маршрутами (в нашем случае их нет) либо составными. Альтернативные составные маршруты уже получены при построении функций $h(p_1^1, p_3), h(p_1^1, p_7), h(p_1^1, p_6)$:

$$m(p_{2}^{1}, p_{3}, p_{1}^{1}) =$$

$$= (p_{2}^{1}, \bar{s}_{2}, \bar{s}_{4}, s_{6}, (p_{3}^{1}, p_{3}^{2}), s_{5}, s_{3}, \bar{s}_{1}, p_{1}^{1}),$$

$$m(p_{2}^{1}, p_{7}, p_{1}^{1}) =$$

$$= (p_{2}^{1}, \bar{s}_{2}, \bar{s}_{4}, \bar{s}_{6}, s_{8}, (p_{7}^{2}, p_{7}^{1}), s_{5}, s_{3}, \bar{s}_{1}, p_{1}^{1}),$$

$$m(p_{2}^{1}, p_{6}, p_{1}^{1}) =$$

$$= (\bar{s}_{2}, \bar{s}_{4}, \bar{s}_{6}, \bar{s}_{8}^{1}, (p_{6}^{2}, p_{6}^{1}), s_{5}, s_{3}, \bar{s}_{1}, p_{1}^{1}).$$

4. АНАЛИЗ МАРШРУТОВ

Покажем, как логические отношения между маршрутами можно определить при представлении структуры железнодорожной станции в виде системы логических функций.

В первом приближении все множество маршрутов можно разделить на совместимые, несовместимые и враждебные маршруты.

Совместимые маршруты — маршруты, по которым поезда могут двигаться одновременно и независимо друг от друга. Это маршруты, которые не имеют общих входных, выходных позиций, а также общих стрелок. Например, маршруты (см. рис. 2)

$$m(p_2^1, p_1^1) = (p_2^1, s_2, s_1, p_1^1),$$

 $m(p_4^1, p_3^1) = (p_4^1, s_4, s_6, p_3^1)$

не имеют общих входных, выходных позиций, а также общих стрелок.

Вывод. Конъюнкции совместимых маршрутов не должны иметь общих переменных.

Несовместимые маршруты — маршруты, которые не могут выполняться одновременно. Пара



маршрутов, отличающихся положением хотя бы одной стрелки, не могут выполняться одновременно, т. е. одновременное движение поездов по несовместимым маршрутам физически невозможно.

Например, конъюнкции ДНФ функции проходимости $f(p_i)$ для всех $p_i \in P$ отличаются друг от друга значением, по крайней мере, одной переменной типа s.

Вывод. Конъюнкции несовместимых маршрутов ортогональны, т. е. их произведение равно нулю.

Враждебные маршруты — пара маршрутов, по которым одновременное движение поездов может привести к аварии (столкновению). Приведем примеры враждебных маршрутов, которые могут быть выявлены путем простого анализа функций проходимости.

1. Пара маршрутов враждебны друг другу, если они имеют одну и ту же конечную позицию. Например, маршруты $m(p_1^1,\ p_3^2)$ и $m(p_2^1,\ p_3^1)$ враждебны, так как они заканчиваются в одной позиции — p_3 .

Вывод. Конъюнкции ДНФ функции проходимости $f(p_i^1)$ представляют маршруты попарно враждебные маршрутам, представленными конъюнкциями ДНФ функции проходимости $f(p_i^2)$ для всех i.

2. Маршруты противоположного направления (встречное движение поездов). Например: конъюнкции $y(p_1^1)s_1^1s_2^2x(p_2^1)$ и $y(p_2^1)s_2^1s_1^2x(p_1^1)$ описывают встречные маршруты. При разрешающих сигналах $x(p_2^1)=1$ и $x(p_1^1)=1$ может иметь место встречное движение поездов.

Вывод. Каждый простой маршрут и ему обратный (реверсивный) являются враждебными маршрутами.

3. Маршруты передвижений по пересекающимся и сливающимся путям. Введение дополнительных позиций в логическую модель в точки пересечения путей сводит этот случай к случаю 1.

ЗАКЛЮЧЕНИЕ

Разработаны логические модели (в виде логических функций) компонентов железнодорожных станций: железнодорожной стрелки, последовательности стрелок, станционных маршрутов поезда (прибытия, отправления, передвижения). Структура железнодорожной станции представлена в виде системы логических функций, на основе которой находится общее множество маршрутов станции. Определены логические отношения между

различными типами маршрутов: совместимых, несовместимых, враждебных.

Полученные результаты могут быть полезны при создании систем поддержки принятия решений станционными диспетчерами.

Автор признателен д-ру техн. наук О.П. Кузнецову, конструктивная и доброжелательная критика которого способствовала улучшению статьи.

ЛИТЕРАТУРА

- 1. *Левин Д.Ю*. Диспетчерские центры и технология управления перевозочным процессом. М.: Маршрут, 2005.
- Cassandras C.G., Lafortune S. Introduction to discrete event systems. — Springer Science + Business Media, LLC, 2008. — 781 p.
- 3. *Амбарцумян А.А.* Моделирование и синтез супервизорного управления на сетях Петри для рассредоточенных объектов // Автоматика и телемеханика. 2011. Ч. 1.: № 8. С. 151—169; Ч. 2.: № 9. С. 173—189.
- 4. Амбарцумян А.А. Потехин А.И. Групповое управление в дискретно-событийных системах // Проблемы управления. 2012. № 5. С. 46—53.
- 5. *Durmuş M.S.*, *Söylemez M.T.* Automation Petri Net Based Railway Interlocking and Signalization Design // International Symposium on Innovations in Intelligent SysTems and Applications, INISTA'09, Karadeniz Technical University, Trabzon, Turkey, 29 June 01 July, 2009.
- Ordonez F., Leachman R., Mural P. Strategies for effective rail track capacity usage / Final Report METRANS Project January 16, 2010. University of Southern California1. Automation Petri Net Based Railway Interlocking and Signalization Design.
- 7. Потехин А.И., Браништов С.А., Кузнецов С.К. Дискретно-событийные модели железнодорожной сети // Проблемы управления. 2014. № 1. С. 73—81.
- Амбарцумян А.А. Дискретно-событийное моделирование в технических системах. Труды и пленарные доклады участников конференции УКИ'12 / Москва, ИПУ РАН, 2012. 1 электрон. опт. диск (CD-ROM). ISBN 978-5-91450-100-3. С. 496—522.
- Isaksson-Lutteman G. Future Train Traffic Control: Development and deployment of new principles and systems in train traffic control / Division of Visual Information and Interaction Department of Information, Technology Uppsala University, 2012.
- D'Ariano A., Corman F., Hansen I.A. Railway traffic optimization by advanced scheduling and rerouting algorithms / Department of Transport & Planning, Delft University of Technology. Delft, The Netherlands, 2010.
- Kecman P., Rob M.P., Goverde R.M.P., van den Boom Ton J.J. A Model-Predictive Control Framework for Railway Traffic Management, 2010.
- 12. *Кузнецов О.П.* Дискретная математика для инженера. СПб.: Лань, 2009.
- Закревский А.Д, Поттосин Ю.В., Черемисинова Л.Д. Логические основы проектирования дискретных устройств. М.: Физматлит, 2007. 589 с.

Статья представлена κ публикации членом редколлегии $O.\Pi.$ Кузнецовым.

Потехин Анатолий Иванович — канд. техн. наук, вед. науч. сотрудник, Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, ⋈ apot@ipu.ru.



CONTENTS & ABSTRACTS

COLLECTIVE INTELLIGENCE TECHNOLOGIES......2 Slavin B.D.

A review is presented of the literature on collective intelligence technologies. Shown is that they differ significantly from crowd sourcing technologies and can be presented as knowledge era sourcing. A description of collective intelligence attributes is suggested that can be used for the design of enterprise knowledge management systems and expert networks. Briefly described are the methods of collective intelligence modeling and measurement. The theory and practice development necessity of collective intelligence technologies is shown.

Keywords: collective intelligence, crowdsourcing, IQ, expert networks, social networks.

Shvedov A.S.

Pointed out is that in classical linear regression model the residuals are assumed to be normally distributed with zero average and standard deviation. However, the real data usually do not satisfy the classical model assumptions. At the same time, even a single outlier can influence significantly on regression parameters estimation. One of the robust regression methods with high breakdown point is the method of least trimmed squares. The new proof of the breakdown point estimation theorem is given, being much more simple that the classic proof.

Keywords: robust regression, least trimmed squares estimator, breakdown point.

Burkov V.N., Korgin N.A., Novikov D.A.

Theoretical problems connected with analysis of integrated control mechanisms for the organizational-technical systems (OTS) are discussed. These systems existing classification extension is offered, arising with transition from separate control mechanisms to integrated ones. A detailed consideration is given of two key problems: the integrated mechanism' efficiency and strategy-proofness assessment and the possibility of its decomposition to assess separate elementary mechanisms.

Keywords: control science, organizational-technical system, systems analysis, strategy-proofness.

The model is suggested of the knowledge intensive firm human capital economical management. Shown is that using the basic principles of enterprise managerial accounting significantly increases the practical usability of the model. Human capital cost optimization problem is solved for the concrete exemplar firm. The example demonstrates the tasks of headcount control and employee development under conditions of irregular and uncertain demand of human resources from core business side.

Keywords: human capital management, knowledge intensive firm, employee lifecycle, stochastic model.

Gribanova E.B.

It is noted that the inverse problems arise from the need in finding optimal ways of companies' issues solving, answering the question «how to make this so, that?..» The article suggests options of the economy inverse problems solving using modified method of inverse calculations. Various types of limitations on function arguments are considered. The possibility is shown to apply modified method of inverse calculations for two arguments function optimization on given interval.

Pointed out is that the obtained results could be useful in the systems of managerial decisions making support.

Keywords: inverse computations, inverse problem, restriction, optimization, management decisions.

TIME SERIES SEGMENTATION AND HASHING IN THE STOCK MARKET PREDICTION PROBLEMS. . . . 41

Spiro A.G., Goldovskaya M.D., Kiseleva N.E., Pokrovskaya I.V.

It is suggested to associate each analyzed time series of exchange-traded asset price (TS-P) with time series of hash codes (TS-HK) that will show the price rising or falling for each element of TS-P. In this case the hash codes are integers, and their sequence allows similar (typical) groups of TS-P highlighting in exchange-traded asset price changing dynamics. Described are the procedures of the original time series conversion and of the corresponding hash codes determination. The basic properties of hash codes sequences are formulated. The methodology of trajectory analysis and the asset exchange rate forecasting is suggested, using segmentation and hashing data.

Keywords: stock market, the procedure of a sliding window, the hash codes, time series segmentation, typical segments, the growth/falling forecasting of stock quotes.

Mikrin E.A., Somov S.K.

The article examines online information backup issues in distributed data processing systems (DDPS) operating within computer networks, in order to improve their operation safety. The features of using three online backup basic strategies in DDPS are analyzed. Problems are formulated of optimal by different criteria DDPS online backup distribution over computing network nodes, and the methods of solving them are proposed. Problems and challenges of data arrays backing up optimization, taking into account the relationship between them, are considered, as well as data backup in distributed databases. Analyzed is the requests traffic impact on DDPS backup distribution.

Keywords: distributed data processing systems, computer networks, backup arrays of data, the optimization problem of the data redundancy, data backup strategy.

Zack Yu.A.

Statement is given and mathematical properties are derived for the problem of constructing acceptable and optimal routing assignments for one or several vehicle types in case of given set of constraints on cargo receiving points and delivery time, and of simultaneous two or more tasks performance impossibility. Based on derived properties of acceptable and optimal sequences, the tasks solution algorithms are developed using modified sequential optimization methods, illustrated by numerical examples. The results of computing experiments are given.

Keywords: sequence of assignments that are valid and optimal routes, estimates of the partial plans, modified sequential optimization algorithms.

LOGICAL MODEL OF RAILWAY STATION OBJECTS 72 Potekhin A.I.

Problems are considered of creating the train dispatching systems, able to predict potential conflict situations. For this purpose a formal method is developed of railway objects (stations, spans) structure representation in the form of a logical functions system, based on which is the number of routes. Logical relations between the routes (compatible, incompatible, conflicting) are determined.

Keywords: logical model, railway station, railroad switch, the trains route, compatible, conflicting routes.

XU

МЕТОДЫ ПРОГНОЗИРОВАНИЯ КРИТИЧЕСКИХ СОБЫТИЙ НА ЭТАПАХ ВЗЛЕТА И ПОСАДКИ САМОЛЕТА

На этапах взлета и посадки воздушных судов происходит более 50% летных происшествий.

Предпосылки неправильных или ошибочных действий экипажа заключаются в высоких психологических нагрузках и необходимости принятия ответственных решений на основе субъективной оценки ситуации.

В целях повышения безопасности воздушных перевозок в ИПУ РАН разработаны алгоритмические и программные средства информационной поддержки экипажа, в том числе:

- алгоритмы оценки текущего и прогнозирования будущего движения воздушного судна по взлётно-посадочной полосе:
- оповещения пилота о наступлении критических событий на взлетно-посадочных этапах полета;
- рекомендации по выполнению или прерыванию типовых полетных процедур.

Все разработки базируются на энергетическом подходе к управлению движением

Математическая формулировка подхода -- УРАВНЕНИЕ БАЛАНСА ЭНЕРГИЙ:

ЭНЕРГИЯ САМОЛЕТА

ВИГАТЕЛЯ

АЭРОДИНАМИКА

BETEP

силы торможения



Прогнозирование безопасного взлета при наличии препятствий



Алгоритм прогнозирования дальности до начала подъема носового колеса

 $D_{\nu} = 0.5((V_{-})^2 - V(t)^2)/gn_{\nu}(t)$

Прогноз дальности до точки принятия решения (ТПР)

$$D_{\text{TUP}}(t) = \left(g\left(H_{\text{input}} - h(t)\right) + 0.5\left(V_2^2 - V(t)^2\right)\right) / gn_x(t) - L_{\text{input}}$$

 $D_{\text{ross}} = 0.5 \left(V(t)^2 - \epsilon^2\right) / g n_x(t)$

Этап торможения

ПРОГНОЗИРОВАНИЕ БЕЗОПАСНОГО ТОРМОЖЕНИЯ



HX FTYCK[2]

Прогноз дальности до точки останова

Коррекция прогноза

 $D_{\text{TORM}} = Q_{\text{TOP}} D_{\text{TORM}} Q_{\text{top}} - Q_{\text{top}} (J_i, k_{\text{tops}}, V)$

Окно настроек моделирующего стенда Serveral point states CMIANAMINE Off. to Off. till with Off. son Crist Avenue

Education (1990)

ED pin 25 | pag 27 | Bar 071

E1 pin 13 | pag 175 | Bar 075 Ken 25 ER 1.05 Stempts martin Martin K1 | 0.50 1.05 V min. 226 ye water ou typoters 20 75 30 45 10 10 10 100 105 Era | 0303040405050454

minut Pennel

for t

выход

Статистические испытания алгоритмов прогнозирования на всем множестве взлетно-посадочных режимов показали их высокую достоверность.

Институт проблем управления им. В.А. Трапезникова РАН

Профсоюзная ул., д. 65, Москва, 117997 anshev@ipu.ru
 (495) 334-90-39