

На правах рукописи
УДК 519.218+519.857.3

Коновалов Михаил Григорьевич

**МОДЕЛИ И ТЕХНОЛОГИИ АДАПТИВНОЙ ОБРАБОТКИ
ИНФОРМАЦИИ ДЛЯ ЧАСТИЧНО НАБЛЮДАЕМЫХ СИСТЕМ**

05.13.17 – теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
доктора технических наук

Москва 2008

Работа выполнена в Институте проблем информатики РАН

Официальные оппоненты:

член-корреспондент РАН, доктор физико-математических наук, профессор
Флеров Юрий Арсениевич

заслуженный деятель науки РФ, доктор технических наук, профессор
Оганян Герман Арташесович

заслуженный деятель науки РФ, доктор технических наук, профессор
Синицин Игорь Николаевич

Ведущая организация:

Московский государственный технический университет им. Н. Э. Баумана

Защита состоится « ___ » _____ 2008 года в ___ часов на заседании диссертационного совета Д 002.073.01 при Институте проблем информатики РАН по адресу: 119333, Москва, ул. Вавилова, 44, корп. 2.

С диссертацией можно ознакомиться в библиотеке Института проблем информатики РАН

Отзывы в двух экземплярах, заверенных печатью, просим направлять по адресу: 119333, Москва, ул. Вавилова, 44, корп. 2, диссертационный совет.

Автореферат разослан « ___ » « _____ » 2008 года.

Ученый секретарь
диссертационного совета Д002.073.01,
доктор технических наук, профессор

С. Н. Гринченко

ОБЩАЯ ХАРАКТЕРИСТИКА ДИССЕРТАЦИИ

Актуальность темы

В различных областях науки и техники имеются многочисленные задачи, для решения которых необходимо анализировать, оптимизировать, реализовывать взаимодействие субъекта с объектом в условиях недостаточной информации. Наибольшее количество подобного рода задач возникает в информационных и телекоммуникационных системах, автоматизированных производственных процессах, робототехнике, то есть в тех сферах, которые в наибольшей степени связаны с компьютерной обработкой информации.

Неполнота информации имеет двойное качество. Во-первых, это (частичное) отсутствие априорной информации, даже на уровне представления о структуре объекта, имеющего, как правило, стохастическую природу. Во-вторых, ограниченная возможность наблюдения объекта и его идентификации. В предельном случае субъекту заранее известно лишь множество, из которого можно выбирать воздействия на объект, а неполнота наблюдений в неблагоприятных случаях означает, что субъект может лишь оценивать отклики объекта с точки зрения своего предпочтения. В подобных ситуациях первостепенное значение приобретает умение воспользоваться доступной информацией об объекте, в том числе, и главным образом, приобретенной в ходе взаимодействия с ним. Умение субъекта пользоваться для достижения цели информацией, поступающей в процессе функционирования, связано с такими понятиями, как «самоорганизация», «приспособление», «адаптация». Диссертационная работа посвящена преимущественно синтезу, анализу и применению адаптивных стратегий обработки информации.

Основополагающие идеи теории адаптации были заложены в 50-х годах прошлого века^{1,2}, а становление теории и ее развитие до конца 80-х годов проходило во многом благодаря усилиям отечественных исследователей³. С начала 90-х годов и по настоящее время это направление переживает большой подъем, а число публикаций исчисляется многими сотнями. В частности, выделилась и получила широкое распространение новая ветвь, тесно связанная с искусственным интеллектом и его разнообразными приложениями⁴.

Центральное место в теории адаптации занимает проблематика частично наблюдаемого марковского процесса принятия решений. Традиционный подход, основанный на динамическом программировании, дал результаты, применимые в адаптивном варианте^{5,6}. «Марковские процессы являются в настоящее время

¹ H. Robbins, Monro S. A stochastic approximation method // Ann. Math. Stat. – V. 22. – 1951. – P. 400–407

² H. Robbins. A sequential decision problem with a finite memory // Proc. Nat. Acad. Sci. USA. – V. 42, No. 3. – 1956. – P. 920–923.

³ V. G. Sragovich. Mathematical Theory of Adaptive Control. – Singapore: World Scientific, 2006.

⁴ R. Sutton, A. Barto. Reinforcement learning. – MIT Press, 2000.

⁵ D. P. Bertsekas. Dynamic programming and optimal control, V. 1, 2. – Belmont: Athena Scientific, 2001.

математическим фундаментом для многих работ в области активного обучения (reinforcement learning), теории принятия решений, поиска информации, распознавания речи, активного зрительного восприятия, навигации роботов»⁷. В то же время «одна из основных проблем в приложении теории марковского процесса принятия решений заключается в необходимости рассматривать не слишком большое пространство состояний»⁸. Как подчеркивают многие авторы, и как свидетельствует большой поток публикаций, необходим дальнейший прогресс в этой области, имеющей неоспоримое прикладное значение.

Развиваемое в диссертационной работе адаптивное направление лежит в русле фундаментальных исследований, как адаптивного, так и не адаптивного характера, в области стохастических систем и стохастических информационных технологий^{9,10}. Стремительный прогресс в области информационных и телекоммуникационных технологий позволил поставить на гораздо более реальную почву вопрос о практической реализации стохастических адаптивных алгоритмов, которые принципиально связаны с быстрой обработкой и передачей больших объемов оперативной информации.

Из вышесказанного следует, что тема диссертационной работы находится в одной из актуальных и плодотворных областей современной науки.

Цель и задачи исследования

Целью диссертационной работы является *разработка теоретических основ анализа, синтеза и применения стратегий адаптивной обработки информации и принятия решений в условиях неполного наблюдения*. При проведении исследований были поставлены следующие основные задачи.

1. Разработка теоретических основ синтеза адаптивных стратегий в общих моделях частично наблюдаемых стохастических объектов.

2. Изучение класса частично наблюдаемых управляемых однородных марковских цепей. Разработка для этого класса простых, эффективных и универсальных адаптивных алгоритмов, позволяющих решать задачи большой размерности.

3. Изучение принципиальной возможности и оценка границ применения теории в различных областях приложений. Разработка общей методологии применения адаптивных методов обработки информации.

4. Реализация комплекса действий, связанных с синтезом информационных технологий для моделирования, адаптивной обработки информации и принятия решений на примерах крупных прикладных проблем.

⁶ H. S. Chang, M. C. Fu, J. Hu, S. I. Marcus. Simulation-based algorithms for Markov decision Processes. – London: Springer, 2007.

⁷ S. Mahadevan. Spatiotemporal abstraction of stochastic sequential processes // Lecture Notes in Computer Science, 2371. – 2002.

⁸ T. Belker, M. Beetz, A.B. Cremers. Learning action models for the improved execution of navigation plans // Robotics and Autonomous Systems, 38. – 2002. – P. 137–148.

⁹ В.С. Пугачев, И. Н. Сеницын. Теория стохастических систем. – М.: Изд. Логос. 2004.

¹⁰ Unsupservised adaptive filtering. V. 1, 2. Edited by S. Haykin. – New York: John Willey & Sons, Inc, 2000.

Методы исследования

Основным аппаратом для формулировки и изучения теоретических вопросов является математическая теория адаптации, которая в значительной мере опирается на теорию вероятностей, стохастический анализ и теорию случайных процессов. Широко используется теория счетных цепей Маркова. В качестве основных моделей объектов выступают классы управляемых случайных последовательностей общего вида с дискретным временем. Для анализа градиентных алгоритмов привлекается математическое программирование. Формальное построение имитационной модели осуществляется с помощью техники параллельных процессов. Компьютерные программные системы выполнены на объектно-ориентированном языке Delphi.

Основные положения, которые выносятся на защиту

1. Теоретические основы синтеза и необходимые условия существования адаптивных стратегий обработки информации для широких классов частично наблюдаемых стохастических объектов.

2. Теоретические основы анализа и технология синтеза градиентных адаптивных стратегий обработки информации для частично наблюдаемого марковского процесса принятия решений.

3. Результаты анализа областей приложения и методология практического применения теории адаптивной обработки информации.

4. Синтез информационных технологий моделирования и адаптивного принятия решений для сети с коммутацией каналов и для распределенной вычислительной среды.

Все заявленные результаты получены лично автором.

Научная новизна

В области теоретических основ адаптивной обработки информации и принятия решения для частично наблюдаемых систем получены следующие новые результаты:

- дано определение класса регенерируемых объектов, который включает частично наблюдаемые управляемые марковские цепи, и доказана теорема существования равномерно ε -оптимальной детерминированной стратегии конечной глубины для этого класса;
- построена оригинальная конструкция адаптивной стратегии перебора и доказана теорема о классе объектов, Σ -управляемых с помощью стратегии перебора;
- доказано, что модификации стратегии перебора являются адаптивными по отношению к классу регенерируемых объектов, по отношению к классу частично наблюдаемых марковских цепей, по отношению к классу частично наблюдаемых графов;
- проведен анализ функции предельного среднего дохода в задаче марковского процесса принятия решений со счетным множеством состояний и получены явные формулы градиента целевой функции для случая полного наблюдения и для случая частичного наблюдения и распределенного характера принятия решений;

- осуществлен синтез и анализ эрланговской модели для сети с коммутацией каналов с адаптивной градиентной стратегией минимизации отказов.

В области применения адаптивных методов новой является методология моделирования и оперативной адаптивной обработки информации для широкого класса стохастических систем с неполным информационным описанием и неполным наблюдением на основе оригинальных адаптивных стратегий.

Практическая ценность и реализация результатов

1) В широком плане практическая значимость результатов диссертации заключается

- в анализе и обосновании областей и примеров приложения теории;
- в технологии синтеза универсальных и эффективных адаптивных алгоритмов для частично наблюдаемых дискретных стохастических объектов;
- в методологии синтеза информационных технологий моделирования и оперативной адаптивной обработки информации.

2) В ходе выполнения НИР «Модель-Т» по разработке модели телефонной сети (совместно с ЗАО «Телесофт–Россия» по заказу ОАО Ростелеком) осуществлен синтез информационных технологий моделирования и адаптивного принятия решений для сети с коммутацией каналов, в том числе разработаны:

- модель и алгоритмы восстановления матрицы тяготения по результатам сетеметрии;
- имитационная модель трафика и принятия оперативных решений на языке параллельных процессов;
- блок адаптивной коррекции обобщенных маршрутных таблиц и оптимизации дополнительных управляющих воздействий.

Разработанные модели и алгоритмы были реализованы в виде программной системы, которая явилась составной частью блока управления трафиком междугородной телефонной сети РФ.

3) При разработке автоматизированных систем управления сетями связи в ходе ОКР «Система-АТМ» и «Север» (ОАО «Интелтех») использованы:

- методология применения алгоритмов оптимизации управления сетевыми ресурсами на базе теории адаптивного управления частично наблюдаемыми марковскими процессами;
- результаты аналитико-имитационного моделирования с использованием адаптивных стратегий оперативного управления для телекоммуникационных сетей.

4) При разработке магистрального широкополосного аппаратного IP-шифратора «Заслон» (ЗАО «Голлард») для достижения оптимальных системотехнических и алгоритмических решений с целью максимальной совместимости с различными приложениями реального времени и различным сетевым оборудованием использованы:

- методология синтеза информационных технологий для моделирования и оптимизации параметров сложных стохастических систем с помощью применения адаптивных алгоритмов в режиме off-line.
- результаты численного моделирования процесса функционирования изделия с использованием адаптивных стратегий оптимизации.

5) Осуществлен синтез информационной технологии моделирования системы распределенных вычислений, в том числе разработаны:

- имитационная модель коллективного взаимодействия взаимно удаленных потребителей и распределенных вычислительных ресурсов;
 - алгоритмы адаптивной обработки информации для оперативного распределения ресурсов при неполной информации;
- Проведено экспериментальное исследование свойств адаптивных алгоритмов в модели системы распределенных вычислений.

Апробация

Результаты исследований и разработок, представленных в диссертации, докладывались на Всесоюзной конференции «Теория адаптивных систем и ее применение» (Ленинград, 1983), на 4-ой международной вильнюсской конференции по теории вероятностей и математической статистике (1985), на XLVII научной сессии, посвященной Дню радио (Москва, 1992), на Международной конференции по информационным сетям и системам ICINAS-98 (С.-Петербург, 1998), на 3-м Всероссийском симпозиуме по прикладной и промышленной математике. (Ростов-на-Дону, 2002), на 3-й Московской международной конференции по исследованию операций (ORM2001), на Международных научно-технических конференциях «Интеллектуальные системы» (IEEE AIS'03) и «Интеллектуальные САПР» (СФВ-2003), на семинаре «Seminar on Stability Problems for Stochastic Models» (Pamplona, 2003), на 8-ой международной конференции «Проблемы функционирования информационных сетей» (Иссык-Куль, 2004), на Международной научно-практической конференции «Информационные технологии и информационная безопасность в науке, технике и образовании ИНФОТЕХ-2004» (Севастополь), на семинаре «XXV International Seminar on Stability Problems for Stochastic Models» (Salerno, 2005), на 2-ой международной конференции «Распределенные вычисления и Грид-технологии в науке и образовании» (Дубна, 2006), на Международном симпозиуме «Информационные технологии и общество» (Тель-Авив, 2007), на семинарах в Вычислительном центре им. А. А. Дородницына РАН, на семинарах в Институте проблем информатики РАН, на семинаре в Институте проблем управления РАН.

Исследования по теме диссертации были поддержаны 7 грантами РФФИ.

Публикации

По теме диссертации имеется 40 публикаций, из них: 9 – в изданиях из перечня ВАК, 1 – монография, 4 – свидетельства о регистрации программ для ЭВМ, 17 – тезисы докладов. Общий объем публикаций – более 25 п. л.

Структура и объем работы

Диссертация состоит из введения, шести глав, заключения и приложения. Основная часть работы изложена на 319 страницах, приложение – на 26 страницах. Работа содержит 46 рисунков и 13 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во **Введении** дана общая характеристика диссертации, ее предмета, цели и задач. Сформулированы основные признаки, которыми отличается изучаемая проблема:

- имеется процесс взаимодействия с объектом произвольной (стохастической) природы, в котором воздействия на объект чередуются с получениями откликов (наблюдений);
- характеристики процесса и возможности наблюдения являются неполными;
- процесс взаимодействия сопровождается синхронной с основным процессом, наблюдаемой последовательностью «доходов»;
- цель взаимодействия – (условная) максимизация предельного среднего дохода;
- процесс чередования пар сигналов «действие – наблюдение» продолжается неограниченно долго (с практической точки зрения – «достаточно долго»).

В этих условиях требуется определить для взаимодействующего субъекта такую стратегию анализа и обработки поступающей информации, которая позволяла бы ему постепенно «обучаться» совершать «правильные» действия и достигать цели, максимально доступной в тех условиях неопределенности, в которых он находится. Разрабатываемые алгоритмы обучения, или стратегии адаптивной обработки информации должны быть устроены как рекуррентные процедуры, осуществляющие целенаправленный бесконечный поиск в конечномерном пространстве на основе статистических оценок. При этом должны выполняться следующие требования:

- **адаптивность** – не предполагается обязательное наличие априорной информации и зависимости от параметров объекта; может использоваться оперативная информация, полученная в ходе взаимодействия с объектом;
- **эффективность** – время адаптации, то есть время до осуществления оптимизации по выбранному критерию качества должно быть небольшим;
- **конструктивность** – реализация алгоритма должна быть простой, сводиться к выполнению на каждом такте сравнительно небольшого числа арифметических операций и требовать сравнительно малой памяти;
- **универсальность** – алгоритмы должны быть применимы к широкому классу объектов разнообразной физической природы, допускать локальное, распределенное и децентрализованное использование.

В **главе 1** на основе анализа большого количества источников сделан обзор значительной части всей сферы приложений, которые классифицированы на рис. 1. В качестве основных областей приложения выделены три: инфотелекоммуникационные системы (**раздел 1.1**), производственные системы (**раздел 1.2**), а также приложения, связанные с моделированием поведения, искусственным интеллектом и робототехникой (**раздел 1.3**).



Рис. 1. Примерная классификация областей приложения адаптивных методов.

Системы компьютерной обработки и передачи информации являются наиболее очевидным направлением применения адаптивных методов. Во многих сферах телекоммуникации они уже с успехом применяются. Например, давно применяется адаптивная маршрутизация. (Приложениям в области маршрутизации для сетей с коммутацией каналов посвящена отдельная глава 5.) Плодотворным приложением является также близкий участок сетевой тематики, связанный с управлением потоками, в частности, важное для беспроводных технологий управление доступом.

Сравнительно новой областью приложения является проектирование систем распределенных вычислений. Необходимость принятия решений возникает при распределении потоков заданий между вычислительными ресурсами, а применение адаптивных методов не нашло пока должного распространения. (Данной проблеме также посвящена отдельная глава 6.) Среди других приложений в области инфотелекоммуникаций указаны поиск информации и организация сайтов в среде Интернет; управление качеством сетевого сервиса; организация поддержки программного обеспечения и некоторые другие.

Следующей важной областью применения адаптивных алгоритмов являются производственные системы: технологические линии, сборочные конвейеры, гибкие автоматизированные производства и пр. Эти системы требуют настройки, балансировки, планирования, диспетчеризации и других оптимизационных решений в области производства и сбыта. Как показывает анализ, такие процессы хорошо укладываются в основную схему, намеченную во введении и изученную далее в главах 2–4. Очень большое количество постановок задач дают также направления, связанные с техническим обслуживанием и управлением запасами.

Первоначальный этап создания и исследования адаптивных алгоритмов в значительной степени проходил под эгидой «моделирования поведения и интеллекта». Это направление вполне сохранилось и достаточно интенсивно развивается в настоящее время. В диссертационной работе приведен обзор моделей поведения, которые предназначены для отработки тех или иных алгоритмов обучения и являются широко разрабатываемой областью применения. К тому же данные исследования воплощаются во вполне осязательных изделиях, которые называются роботами. Приведено несколько иллюстративных примеров из робототехники.

Раздел 1.4 содержит примеры из самых разнообразных областей приложений: управляемые системы массового обслуживания (за которыми также стоят всевозможные применения), примеры из механики, экологии и пр.

В **главе 2** приведены базовые определения и конструкции, необходимые для изучения теоретических основ адаптивной обработки информации при неполном наблюдении. В **разделе 2.1** дано общее определение объекта как дискретной частично наблюдаемой управляемой случайной последовательности. Сформулировано общее понятие стратегии.

В качестве общей модели рассмотрен процесс, имеющий три компоненты: (x_t, y_{t+1}, z_{t+1}) , $t = 0, 1, \dots$, которые называются соответственно *состоянием*,

действием и наблюдением. Эволюция этих компонент определяется с помощью трех последовательностей условных распределений: $\mu = (\mu_0, \mu_1, \dots)$, $\nu = (\nu_1, \dots)$ и $\sigma = (\sigma_1, \dots)$, определяющих соответственно изменение состояний, появление наблюдений и выбор действий. Пара (μ, ν) называется *объектом*, последовательность σ называется *стратегией*, а элементы последовательности σ называются *правилами*.

Наглядно, процесс развивается следующим образом (рис. 2). В начальный момент $t = 0$ состояние x_0 объекта определено согласно распределению μ_0 . В момент $t = 1$ субъект должен выбрать некоторое действие y_1 (он делает это с помощью правила σ_1); в результате субъект может произвести наблюдение z_1 , которое возникает по закону ν_1 . Если процесс достиг предыстории x^{t-1}, y^t, z^t , то состояние объекта x_t возникает с помощью распределения μ_t . Очередное действие y_{t+1} субъект выбирает, руководствуясь правилом σ_{t+1} , а новое наблюдение появляется в силу распределения ν_{t+1} и т. д.

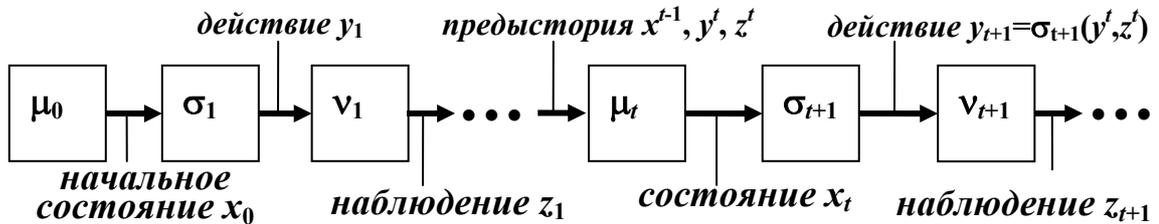


Рис. 2. Процесс взаимодействия субъекта и объекта.

Если в наблюдении можно выделить компоненту, полностью совпадающую с состоянием, то объект называется *полностью наблюдаемым*, а в иных случаях объект называется *частично наблюдаемым*. Предполагается, что в любом случае наблюдение содержит компоненту, которая принимает числовые значения из интервала $(0, 1)$ и называется *доходом*.

Выделяются некоторые типы стратегий, которые важны для анализа поведения субъекта. Так, стратегия называется *детерминированной*, если все ее правила задаются с помощью вырожденных (сосредоточенных в одной точке) условных распределений. Стратегия имеет *глубину* $d \geq 0$, если все ее правила, начиная с момента $d + 1$, существенно зависят только от предыстории глубины d до момента t включительно. Если состояния объекта наблюдаемы, то стратегия, состоящая из правил, у которых выбор очередного действия зависит только от предыдущего состояния, называется *марковской стратегией*. Стратегия глубины d называется *однородной*, если все ее правила, начиная с момента $d + 1$, не зависят явно от времени. Таким образом, по определению, в однородной стратегии глубины d , начиная с момента $d + 1$, применяется одно и то же правило глубины d .

В диссертационной работе изучаются следующие основные объекты (1–3).

1. Одно из центральных мест отведено изучению (*управляемых*) однородных *марковских цепей* со счетным или конечным множеством состояний, конечным множеством действий и с доходами, причем доход определяется последним состоянием процесса и выбранным в этом состоянии действием.

Принципиально различаются случаи полного и частичного наблюдения, каждый из которых является предметом изучения.

2. Если в управляемой однородной марковской цепи при заданном действии переход от одного состояния цепи к другому происходит детерминировано, то говорим об (*управляемом*) *графе с доходами*. Конечное множество состояний играет роль *вершин* графа, находясь в которых можно совершить действие из заданного конечного множества. С переходом по *дуге* вида (i, i_1, k) , который осуществляется, если в вершине i выбрано действие k , связан доход – случайная величина, имеющая распределение, зависящее от i и k . Изучается случай частично наблюдаемого связного графа.

3. Для теоретического анализа проблемы поведения в условиях недостаточной информации представляет интерес (частично наблюдаемый) *регенерируемый объект*, с конечными множествами состояний и действий, который обладает следующими свойствами. На любом этапе взаимодействия можно «организовать» с помощью выбора подходящих действий наступление момента регенерации. В этот момент объект переводится в состояние, «эквивалентное» начальному состоянию. Состояние объекта в текущий момент времени зависит от всей предыстории с момента предыдущей регенерации, причем характер зависимости может быть произвольным, а промежутки между моментами регенерации могут быть сколь угодно длинными. В довершение состояния объекта не наблюдаемы (число их неизвестно), а наблюдаемы только доходы. Введенный объект обобщает понятие частично наблюдаемой связной марковской цепи. Появление его вызвано стремлением указать как можно более широкую конструкцию объекта, с которым можно целенаправленно взаимодействовать, «обучаясь» в процессе взаимодействия. При этом наличие моментов регенерации представляется совершенно необходимым, для того чтобы избежать «непоправимых» ошибок в процессе обучения.

В **разделе 2.2** показана фундаментальная роль детерминированных однородных стратегий конечной глубины. Рассмотрены регенерируемые объекты, для которых определено понятие предельного среднего дохода за стратегию. Доказана **теорема**, которая устанавливает, что для любого указанного объекта существует детерминированная однородная равномерно ε -оптимальная относительно начального распределения стратегия конечной глубины. Это утверждение является аналогом ряда известных результатов, согласно которым структура зависимости правил стратегии от предыстории повторяет структуру зависимости переходных вероятностей объекта от предыстории. В связи с полученным результатом надо подчеркнуть, что на характер зависимости переходных вероятностей для регенерируемых объектов не накладываются ограничения ни по однородности, ни по глубине предыстории, ни по каким-либо условиям эргодичности. Доказанная теорема имеет важные следствия в двух отношениях. Во-первых, множество стратегий, среди которых следует искать равномерно ε -оптимальную стратегию для рассматриваемых объектов, является счетным. Во-вторых, данный класс объектов содержит как частный случай частично наблюдаемые связные марковские цепи.

Глава 3 посвящена изучению метода перебора в теории адаптации. Адаптивной стратегией принято называть стратегию, которая достигает цели для любого объекта из заданного множества объектов и не зависит при этом от переходных вероятностей, определяющих объект. Один из наиболее эффективных приемов построения адаптивной стратегии основан на простой идее «перебора вариантов». В процессе взаимодействия с объектом перебираются возможные варианты поведения. Если какой-то вариант представляется «хорошим», то его применяют достаточно долго, прежде чем перейти к испытанию нового варианта. Поскольку качество различных вариантов можно оценить лишь со случайными ошибками, то процесс перебора никогда не оканчивается, но среднее время применения «хорошего» варианта подавляюще превосходит среднее время использование менее выгодных вариантов. Первая реализация этой идеи – автомат Роббинса², который предназначался для однородных процессов с независимыми значениями. В качестве множества «вариантов» на каждом такте взаимодействия в этом случае выступает (конечное) множество действий. В дальнейшем это направление широко развивалось, преимущественно применительно к процессам с независимыми значениями. В диссертации содержится широкое обобщение данного метода.

Более сложные объекты требуют перебора на все более разнообразных множествах. Например, в случае конечных марковских цепей необходимо перебирать конечное множество всех детерминированных однородных марковских правил. Еще сложнее обстоит дело в случае частично наблюдаемых конечных марковских цепей, для которых, как показано в разделе 2.2, перебирать нужно счетное количество вариантов. Для организации такого перебора необходимо преодолеть следующие принципиальные затруднения. Во-первых, его надо осуществить на бесконечном множестве и, во-вторых, качество того или иного варианта определяется в терминах предельного вероятностного распределения. Любое из этих обстоятельств означает, что поиск наилучшего варианта никогда не может окончиться, и что к любому варианту для его испытания необходимо возвращаться неограниченное число раз.

В **разделах 3.1, 3.2** исходной позицией является некоторое заданное счетное множество стратегий. Рассмотрен класс всех объектов, для которых любая стратегия из указанного счетного множества является допустимой, и при этом выполнены некоторые условия эргодичности. Построена адаптивная *стратегия перебора*, и доказана **теорема** о том, что для любого объекта из обозначенного класса с помощью этой стратегии можно получить предельный средний доход, равный точной верхней грани доходов по заданному счетному множеству стратегий. Этот результат можно более наглядно перефразировать так. Задано счетное количество «алгоритмов поведения» и имеется объект, свойства которого заранее неизвестны, и к которому можно применить любой из этих алгоритмов. Построена стратегия, которая достигает той же цели, что и наилучший из алгоритмов.

При построении адаптивной стратегии метод перебора реализован следующим образом. Обозначим $\varphi_{t,n} = \frac{1}{n} \sum_{s=t+1}^{t+n} g_s$, где g_t – последовательность доходов, а n – натуральное число, и положим $\theta_{t,n} = \text{Int}((1 - \varphi_{t,n})^{-n})$, где $\text{Int}(\cdot)$ означает целую часть числа. Зададим последовательность марковских моментов с помощью рекуррентных соотношений $\tau_0 = 0$, $\tau_n = \tau_{n-1} + n + \theta_n$, где $\theta_n = \theta_{\tau_{n-1},n}$. Зададим также последовательность случайных величин β_n , принимающих целые положительные значения.

Пусть Σ – заданное счетное множество стратегий. Стратегия перебора $\sigma^* = \sigma^*(\Sigma, \beta)$ определяется формулой $\sigma_t^* = \sum_{n=1}^{\infty} \sigma_t^{(\beta_n)} I_{\{\tau_{n-1} < t \leq \tau_n\}}$, $\sigma^{(i)} \in \Sigma$, где

I – индикаторная функция. Развернутое описание стратегии σ^* выглядит следующим образом. Процесс управления разбивается на *этапы*. Этап с номером n начинается в момент $\tau_{n-1} + 1$ и оканчивается в момент τ_n . В момент, предшествующий началу очередного этапа, определяется номер стратегии, принадлежащей множеству Σ , из которой будут взяты правила для применения на данном этапе. Этот номер равен значению случайной величины β_n . Продолжительность этапа равна $n + \theta_n$ и зависит, следовательно, от номера этапа и от оценки качества применяемой стратегии, полученной в течение первых n тактов.

Интересно отметить, что на каждом этапе стратегии перебора в течение θ_n тактов (с ростом значения n это будет подавляющая часть времени) текущая информация не обрабатывается.

В **разделе 3.3** с помощью метода перебора решена задача построения адаптивной стратегии для класса регенерируемых объектов (доказана соответствующая **теорема**). В **разделе 3.4** доказана аналогичная **теорема** с более сильным результатом для класса однородных конечных связных частично наблюдаемых марковских цепей.

В **разделе 3.5** рассмотрен еще более узкий класс графов с детерминированными переходами между конечным множеством ненаблюдаемых вершин и доходами, связанными с пребыванием в вершинах. Количество вершин неизвестно. Доказана **теорема** о том, что стратегия перебора обеспечивает для любого такого объекта предельный средний доход, равный тому, который можно было бы получить, располагая возможностью полного наблюдения.

Глава 4 посвящена изложению нового направления в развитии известной концепции марковского процесса принятия решений. Предложенный подход основан на анализе целевой функции предельного среднего дохода и на явном аналитическом выражении для градиента этой функции. Сформулирован базовый алгоритм градиентной оптимизации целевой функции. Сделаны важные с точки зрения приложений обобщения на случай неполного наблюдения и на ситуацию, когда взаимодействие с объектом осуществляется рас-

пределенным образом. Важнейшее значение для эффективного применения градиентных алгоритмов имеет способ оценки градиента целевой функции. Если переходные матрицы процесса неизвестны, то градиент не может быть вычислен и необходимы его оценки по результатам наблюдений. В этом случае градиентный алгоритм превращается в адаптивную стратегию.

Градиентный подход в марковском процессе принятия решений появился позже других методов. Первые из известных работ этого направления¹¹ скорее «обозначали желание» двигаться по градиенту, чтобы оптимизировать целевую функцию, и не дали конструктивных алгоритмов. Основная формула для градиента была опубликована в 1989 г.¹², и там же был указан путь ее использования, в том числе в условиях неполного наблюдения и распределенного характера взаимодействия. Начиная примерно с середины 90-х годов, за рубежом появилось много публикаций, связанных с градиентным подходом^{13,14}. Исследования, которые излагаются в диссертации, проводились независимо от этих работ и не имеют среди них аналогов.

В разделе 4.1 рассматривается однородная управляемая связная марковская цепь со счетным множеством состояний X , и конечным множеством действий Y . Вероятность перехода в состояние i при условии, что цепь находится в состоянии j и выбрано действие k обозначается $Q_{ij}^{(k)}$. Одношаговый доход в состоянии i при выборе действия k считается детерминированным и обозначается $G_i^{(k)}$.

Пусть $0 \in Y$. Однородную марковскую стратегию можно отождествить с матрицей S с элементами $S_i^{(k)}$, которые означают вероятность выбора действия $k \in Y \setminus \{0\}$ при условии, что в предыдущий момент система находилась в состоянии i . Для любого $\delta > 0$ такого, что $|Y| \cdot \delta \leq 1$, определим множество стратегий $\mathbf{S}_\delta = \left\{ S : \sum_{k>0} S_i^{(k)} \leq 1 - \delta, S_i^{(l)} \geq \delta, i \in X, l \in Y \setminus \{0\} \right\}$ и положим

$\mathbf{S} = \bigcup_{\delta>0} \mathbf{S}_\delta$. Если $S \in \mathbf{S}$, то матрица $P = P(S)$ с элементами

$$P_{ij} = \sum_{k>0} S_i^{(k)} Q_{ij}^{(k)} + \left(1 - \sum_{k>0} S_i^{(k)} Q_{ij}^{(k)} \right) Q_{ij}^{(0)}$$
 является переходной матрицей

обычной (неуправляемой) марковской цепи с множеством состояний X . Предполагается, что если $S \in \mathbf{S}$, то цепь P является апериодической эргодической с эргодическим классом X , и поэтому предельный средний доход за

¹¹ El-Fattah Y. M. Gradient approach for recursive estimation and control in finite Markov chains // Adv. Appl. Probab., 13. – 1981. – С. 778–803.

¹² Работа [8] в списке литературы, приведенном в конце автореферата.

¹³ Cao X.-R. The Relations Among Potentials, Perturbation Analysis, and Markov Decision Processes // Discrete Event Dynamic Systems: Theory and Applications, 8. – 1998. – P. 71–87.

¹⁴ Marbach P., Tsitsiklis J.N. Approximate gradient methods in policy-space optimization of Markov reward processes // Discrete Event Dynamic Systems: Theory and Applications, 13. – 2003. – P. 111–148.

стратегию S равен $W(S) = \pi\gamma$, где π – вектор-строка предельных вероятностей состояний, а γ – вектор-столбец с компонентами $\gamma_i = G_i^{(0)} + \sum_{k>0} S_i^{(k)}(G_i^{(k)} - G_i^{(0)})$.

Доказана **теорема** об унимодальности функции $W(S)$ на множестве \mathbf{S} .

При некоторых дополнительных предположениях доказан следующий основополагающий результат.

Теорема. Частные производные предельных вероятностей $\pi(S)$, $S \in \mathbf{S}$, конечны и имеют вид $\frac{\partial \pi}{\partial S_a^{(b)}} = \pi_a \sum_{t=0}^{\infty} \kappa_a^{(b)} P^t$, где $\kappa_a^{(b)}$ – вектор-строка с компонентами $\kappa_{aj}^{(b)} = Q_{aj}^{(b)} - Q_{aj}^{(0)}$, $a \in X$, $b \in Y \setminus \{0\}$. Частные производные предельного среднего дохода задаются формулами

$$\frac{\partial W(S)}{\partial S_a^{(b)}} = \pi_a \left(\sum_{t=0}^{\infty} \kappa_a^{(b)} P^t \gamma + G_a^{(b)} - G_a^{(0)} \right).$$

Выражение для частных производных целевой функции допускает ясную интерпретацию. Пусть $H_i^{(k)}(n)$ означает средний доход к моменту n , который получается, если начальное состояние цепи (в момент 0) было i и при этом было выбрано действие k . Тогда с точки зрения критерия максимизации предельного среднего дохода следует увеличить (уменьшить) вероятность выбора в состоянии i действия k за счет уменьшения (увеличения) вероятности выбора действия l , если $H_{kl} = \lim_{n \rightarrow \infty} (H_i^{(k)}(n) - H_i^{(l)}(n)) > 0$ ($H_{kl} < 0$). Подобные соображения открывают путь для построения алгоритмов оптимизации марковского процесса принятия решений.

В **разделе 4.3** установлены некоторые дополнительные свойства целевой функции $W(S)$. В **разделе 4.4** получено другое удобное для применения представление частных производных предельного среднего дохода.

В **разделе 4.5** изучается способ отыскания максимума функции $W(S)$ на замкнутом выпуклом множестве \mathbf{S}_δ , который заключается в построении рекуррентных алгоритмов градиентного типа. Пусть цепь имеет конечное множество состояний. Рассмотрим последовательность $S_{t+1} = \Pi_\delta[S_t + a_t v_t]$, где $S_t \in \mathbf{S}_\delta$, a_t – заданная числовая последовательность, v_t – некоторая случайная последовательность той же размерности, что и S_t , $\Pi_\delta[\cdot]$ – оператор проектирования на множество \mathbf{S}_δ , S_0 – фиксированная точка из множества \mathbf{S}_δ . Последовательность S_t задает стратегию, которую обозначим через σ и назовем *градиентной адаптивной стратегией*. Согласно стратегии σ , если цепь в момент t находится в состоянии i , то выбор управления осуществляется с помощью распределения, содержащегося в i -ой строке матрицы S_t . Это распределение является, вообще говоря, разным в разные моменты времени, поэтому стратегия σ неоднородная. Если последовательность v_t зависит не

только от текущего состояния цепи, но и от предыдущих состояний и, возможно, от предыдущих действий, то стратегия σ не является марковской. Последовательность v_t играет роль оценки градиента. Доказан следующий результат относительно свойств градиентной адаптивной стратегии.

Теорема. Пусть σ – градиентная стратегия и пусть числовая последовательность a_t подчиняется условиям $\sum a_t = \infty$, $\sum a_t^2 < \infty$, а последовательность случайных векторов v_t с измеримыми относительно предыстории компонентами удовлетворяет неравенству $\|v_t - \nabla W(S_t)\| \leq b_t$, почти наверное относительно меры, порожденной стратегией σ . Пусть числовая последовательность b_t такова, что $\sum a_t b_t < \infty$. Тогда почти наверное выполняются равенства $\lim_{t \rightarrow \infty} W(S_t) = \max_{S \in S_\delta} W(S)$, $\lim_{t \rightarrow \infty} d(S_t, S^*) = 0$, где S^* – множество точек максимума функции $W(S)$ на множестве S_δ , а d – расстояние от точки до множества в евклидовой метрике.

В разделе 4.6 классическая модель процесса принятия решений обобщена на ситуацию, которая весьма распространена на практике. Обобщение касается двух направлений: 1) распределенного взаимодействия и 2) отсутствия полной наблюдаемости. Наглядное описание новой схемы в простейшем случае выглядит следующим образом (рис. 3). Предполагается, что в системе имеются две «подсистемы», каждая из которых выбирает свое действие – параметр. Результирующее действие системы, является функцией от этих параметров.

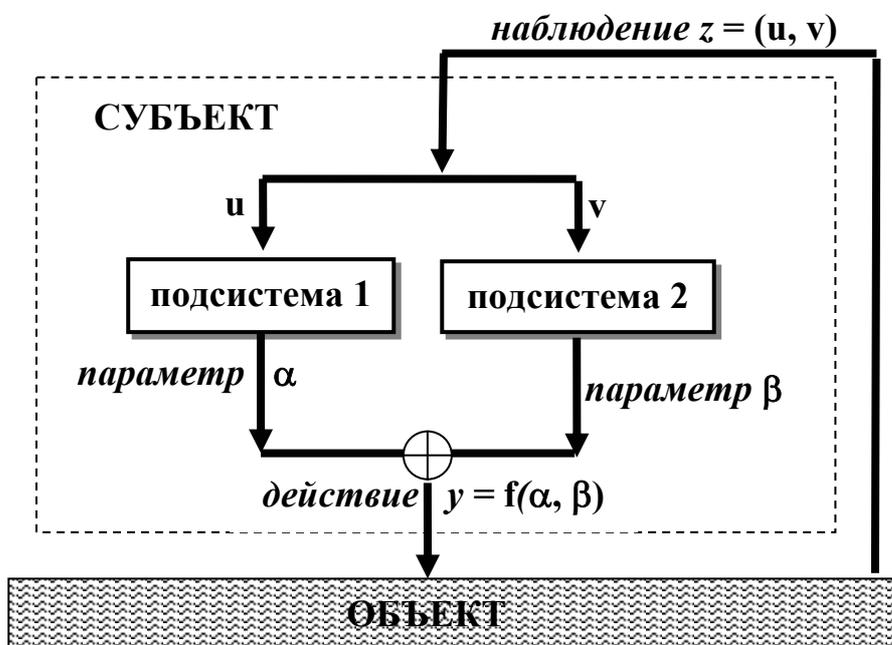


Рис. 3. Распределенное взаимодействие при неполном наблюдении.

Выбор значения параметра внутри каждой подсистемы осуществляется в зависимости от собственных наблюдений с помощью собственных стратегий, которые опираются на эти наблюдения. Получена формула для частных производных целевой функции по компонентам стратегий подсистем, которая

подобно приведенной выше основной формуле, удобна для практической реализации аналога градиентной стратегии

Градиентный подход к оптимизации на марковских цепях оказался возможным благодаря конструктивной формуле для частных производных целевой функции по компонентам стратегии управления, которая переносится также на случай неполного наблюдения состояний цепи. Однако для реализации алгоритма необходимо знать переходные матрицы процесса. Если такой информации нет, то возникает проблема оценивания градиента по результатам наблюдения за траекторией процесса, а алгоритм превращается в адаптивную стратегию. В разделе 4.7 разработаны эффективные методы оценки градиента. Рассмотрим градиентную стратегию, порождающую последовательность марковских правил S_t . Обозначим через $\xi = \xi(i)$, $i \in X$, семейство условно независимых случайных величин, принимающих значения из интервала $(0, 1)$, и положим $\xi_t = \xi(x_t)$, $m_t = \sum_{i \in X} \mathbf{M}\xi(i) \cdot \pi_i(S_t)$, где $\pi(S_t)$ –

предельное распределение, соответствующее марковскому правилу S_t . Отметим, что если $\xi_t = g_t$, то речь идет об оценке предельного среднего дохода $W(S_t)$, соответствующего тому марковскому правилу, которое предписывает алгоритм в момент t . Если же $\xi_t = I_{\{x_t=i\}}$, то оценивается предельная вероятность состояния i , соответствующая марковскому правилу S_t . С учетом того, что величины S_t , изменяются, вообще говоря, в каждый момент, а оценить надо величину, усредненную по предельному распределению, задача построения таких оценок оказывается нетривиальной. Алгоритм оценивания величины m_t строится следующим образом.

Определим последовательности, α_t , ζ_t и β_t следующими формулами:

$$\begin{aligned}\alpha_t &= 1 - (t+1)^{-\alpha}, \quad 0 < \alpha < 1; \\ \zeta_{t+1} &= \alpha_t \beta_t \zeta_t + \xi_t, \quad \zeta_0 = \xi_0; \\ \beta_{t+1} &= \alpha_t \beta_t + 1, \quad \beta_0 = 1.\end{aligned}$$

В качестве оценки величины m_t рассмотрим величину $\mu_t = \zeta_t / \beta_t$.

Теорема. Пусть σ – градиентная стратегия, и пусть выполнено условие $\|S_{t+1} - S_t\| \leq Ct^{-a}$, где $C > 0$, $\max(\alpha, 1/2) < a < 1$. Тогда $|\mu_t - m_t| = O(t^{-b})$ почти наверное относительно меры, порожденной стратегией σ , где $b \in (0, \min\{\alpha, a - \alpha\})$.

В разделе 4.8 приведены результаты вычислительного эксперимента, демонстрирующих то свойство оценок градиента, которое делает адаптивную градиентную стратегию особенно эффективной. В рассмотренном примере значение аргумента S_t (марковского правила) изменяется с постоянной скоростью, что приводит к постоянному изменению значения градиента. Результаты эксперимента показывают (рис. 4), что оценка градиента достаточно хорошо «отслеживает» траекторию точного значения величины $\nabla W(S_t)$. Мас-

штаб по оси времени зависит от размерности объекта, которая, однако, не влияет на качественную картину графиков.

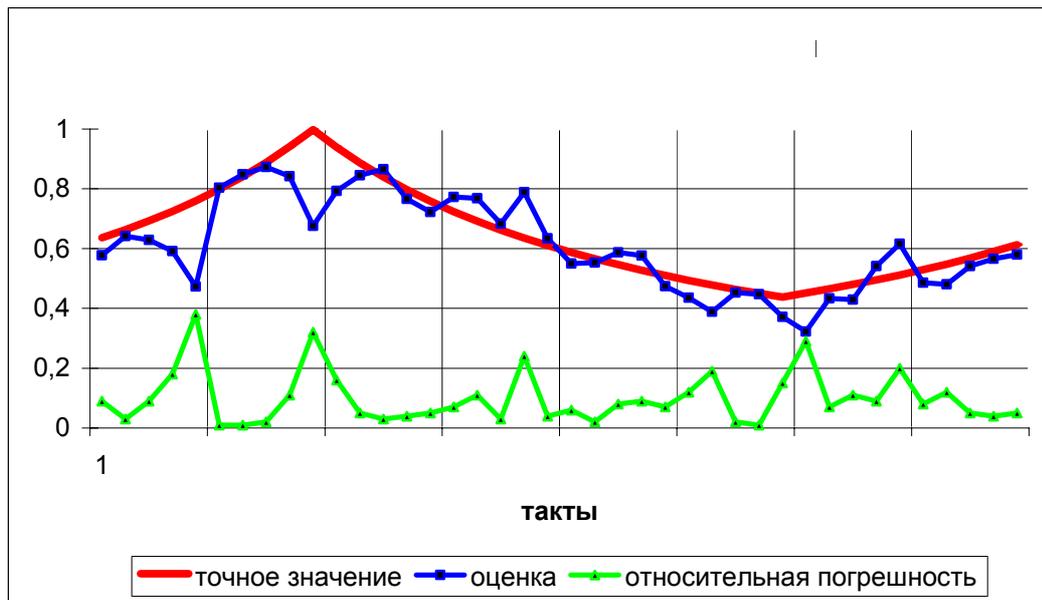


Рис. 4. Оценка градиента целевой функции.

Глава 5 посвящена практическому применению разработанной теории. В ней описана общая методология обработки информации и принятия решений в распределенной системе взаимодействия с частично наблюдаемым объектом, о котором имеется неполная априорная информация. На основе этой методологии осуществляется синтез информационных технологий для оперативного управления трафиком в цифровой сети с коммутацией каналов. Основные положения, которые содержит предлагаемый подход, изложены в **разделе 5.1**. Они иллюстрируются рисунком 5 и заключаются в следующих пунктах.

1. *Выбор* подходящего *масштаба времени* и выделение тактов, то есть указание последовательности моментов, в которые будет происходить взаимодействие с объектом.
2. *Определение информационного обмена* с объектом. Этот пункт включает описание множества воздействий на объект, а также множества доступных на каждом такте наблюдений. Кроме того, необходимо установить, что будет являться одношаговым доходом.
3. *Декомпозиция процесса взаимодействия*, то есть разбиение системы на подсистемы с указанием множества действий и множества наблюдений, доступных для каждой подсистемы на одном такте, а также характера взаимодействия подсистем. На рис. 3 показан пример такой распределенной структуры, в которой отдельные части системы располагают различными компонентами наблюдения и имеют свои наборы параметров, значения которых должны устанавливаться на каждом шаге. «Фактическое» воздействие на объект является результатом решений, принимаемых отдельными подсистемами.
4. Максимально возможное априорное *сужение множества допустимых стратегий* для каждой подсистемы. Эта процедура может осуществляться

с помощью аналитических методов, экспертных оценок, учета технологических особенностей и пр.

5. *Параметризация множеств действий* для каждой подсистемы с помощью конечных наборов конечнозначных параметров.

6. *Построение имитационно-аналитической модели* включает:

- анализ априорной информации об объекте с целью определения структуры и набора параметров взаимодействия;
- создание аналитических методов, позволяющих рассчитывать параметры взаимодействия с объектом;
- написание формальной имитационной модели на специализированном языке параллельных процессов;
- разработку методов идентификации параметров модели;
- построение компьютерной программной системы.



Рис. 5. Схема обработки информации в распределенной системе взаимодействия с частично наблюдаемым объектом при неполной априорной информации; ➡ – информационные потоки.

При выполнении пунктов 1–6 используется максимально доступная априорная информация об объекте.

7. Применение адаптивных стратегий для обработки информации и оптимизации выбора параметров, ответственных за принятие одношаговых решений.

Разработанная методология предполагает два способа реализации: прямое взаимодействие с объектом (режим on-line) и использование имитационно-аналитической модели (режим off-line). Наличие имитационной модели не является обязательным в тех случаях, когда взаимодействие с объектом происходит «достаточно быстро» и «достаточно долго». Но и в этих случаях роль имитационной модели всегда положительна, поскольку появляется возможность исследовать различные сценарии, избегая рисков непоправимых ошибок.

Реализация адаптивных алгоритмов является однотипной для каждой из подсистем, образованных в результате выполнения пункта 3. При этом обработка информации имеет двухуровневую организацию (рис. 6). На верхнем уровне накопленная к моменту t оперативная информация (то есть доступная для данной подсистемы часть наблюдений за объектом) обрабатывается с целью определения сценария, которому удовлетворяет имеющаяся предыстория. Формально это означает отображение f предыстории наблюдений глубины d , $z_{t-d+1}^t = (z_t, z_{t-1}, \dots, z_{t-d+1})$ в конечное множество сценариев $C = \{1, \dots, C\}$. Количество сценариев, отображение f , а также глубина учитываемой предыстории d определяются на предварительных этапах 3 и 4. Каждому сценарию $\tilde{n} \in \tilde{N}$ соответствует адаптивная стратегия A_c (например, стратегия перебора или градиентная стратегия), которая представляет собой последовательность правил, отображающих предысторию доходов в конечное множество значений параметров данной подсистемы. Если на t -м такте

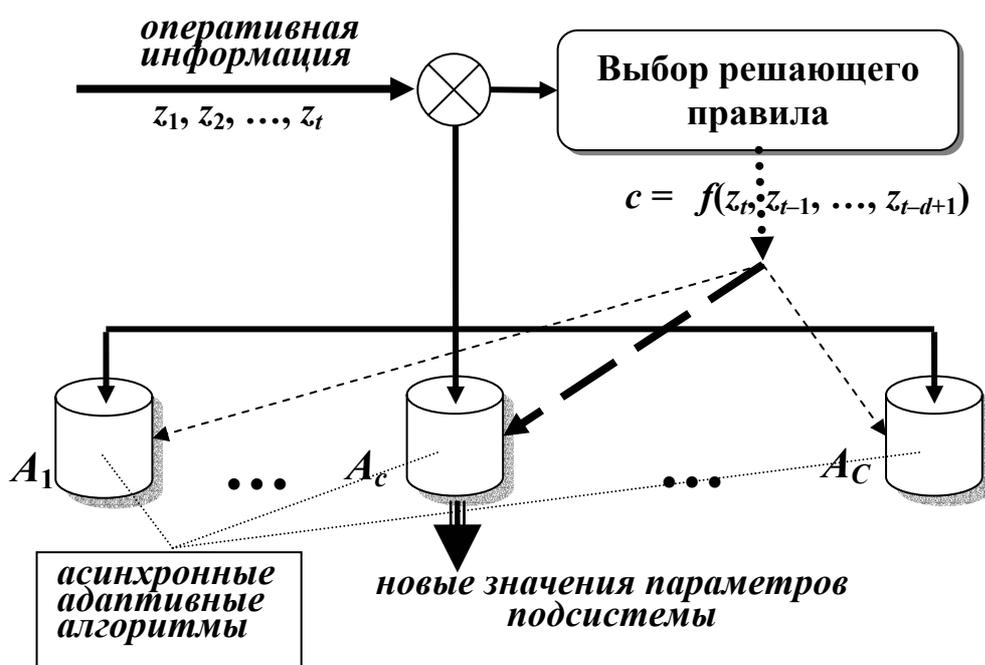


Рис. 6. Двухуровневая реализация адаптивной стратегии обработки информации и принятия решений в подсистеме.

реализовался сценарий $c = c_t$, то *решающим правилом*, которое определяет значение параметров подсистемы на следующем $(t+1)$ -м такте, является правило из стратегии A_c . Таким образом, нижний уровень обработки оперативной информации представляет собой совокупность алгоритмов $\{A_c, c \in C\}$, которые работают, вообще говоря, в асинхронном режиме. Принимаемое на каждом такте решение определяется ровно одним из этих алгоритмов.

В качестве примера реального приложения разработанной теории в диссертации рассмотрена проблема оперативного управления вторичной телефонной сетью в режиме, приближенном к реальному времени, в том числе при возникновении внештатных ситуаций. Рассмотрена многоуровневая сеть коммутации каналов, которая удовлетворяет рекомендациям Международного союза электросвязи¹⁵ и аналогична по функциям и масштабам междугородной сети РФ. На ее примере демонстрируются основные черты методологии и технологии моделирования большой системы и применения адаптивных методов обработки информации и принятия решений.

Все данные относятся к периоду проведения работ (1999–2001 гг.). Материалы соответствуют существовавшей на этот период концепции единой интегрированной автоматизированной системы управления для существующей и перспективной цифровой сети связи АО «Ростелеком» АСУ ЦС. Указанная концепция отвечала всем современным требованиям и учитывала опыт ведущих разработчиков систем управления сетями связи.

Ключевым моментом синтеза информационных технологий для оперативного управления трафиком в цифровой сети с коммутацией каналов является применение разработанных адаптивных алгоритмов для коррекции маршрутных таблиц и оптимизации дополнительных управляющих воздействий. В то же время процесс такого синтеза является многоплановым и включает несколько этапов и составных частей:

- изучение реального объекта и анализ его основных технологических особенностей в контексте поддержки принятия решений по управлению трафиком;
- создание упрощенной математической модели, отражающей важнейшие особенности объекта (без учета ряда специфических деталей), постановку и решение задач оптимизации;
- разработка методов идентификации матрицы тяготения по результатам сетевых измерений;
- формальный синтез объемлющей имитационной модели, включающей такие черты объекта, которые не поддаются аналитическому изучению;
- анализ специфических средств управления трафиком с позиций общей теоретической модели адаптивной обработки информации и применение адаптивных алгоритмов;
- компьютерная реализация комплекса моделей.

¹⁵ Recommendation E. 412 (10/92). Telephone network and ISDN. Quality of service, network management and traffic engineering. Network management controls. – ITU, 1993.

В разделе 5.2 сконцентрировано описание основных аспектов организации трафика в исследуемой системе. Приведена общая характеристика АСУ ЦС ОАО «Ростелеком» и детально перечислены основные объекты сети. Дано описание подсистемы управления вторичной телефонной сетью с указанием ее управляемых объектов. Разобраны функции блока управления трафиком (рис. 7). Анализ показал, что специфика современных телефонных сетей, определяемая как рекомендацией E.412 Международного союза электросвязи, так и реальными параметрами телефонных сетей РФ, не позволяет использовать «напрямую» ни одну из описанных в научной литературе моделей телефонных сетей. Эти модели, как правило, используют ограничительные предположения, далеко не всегда выполняющиеся на практике; в частности, эти предположения не выполняются и в сетях, удовлетворяющих рекомендации E.412. В связи с этим была поставлена задача создания сложной многоуровневой модели, включающей имитационную модель собственно телефонной сети с учетом всей специфики ее деятельности и алгоритмы динамического управления сетью, использующие как имитационную, так и аналитическую модель сети.

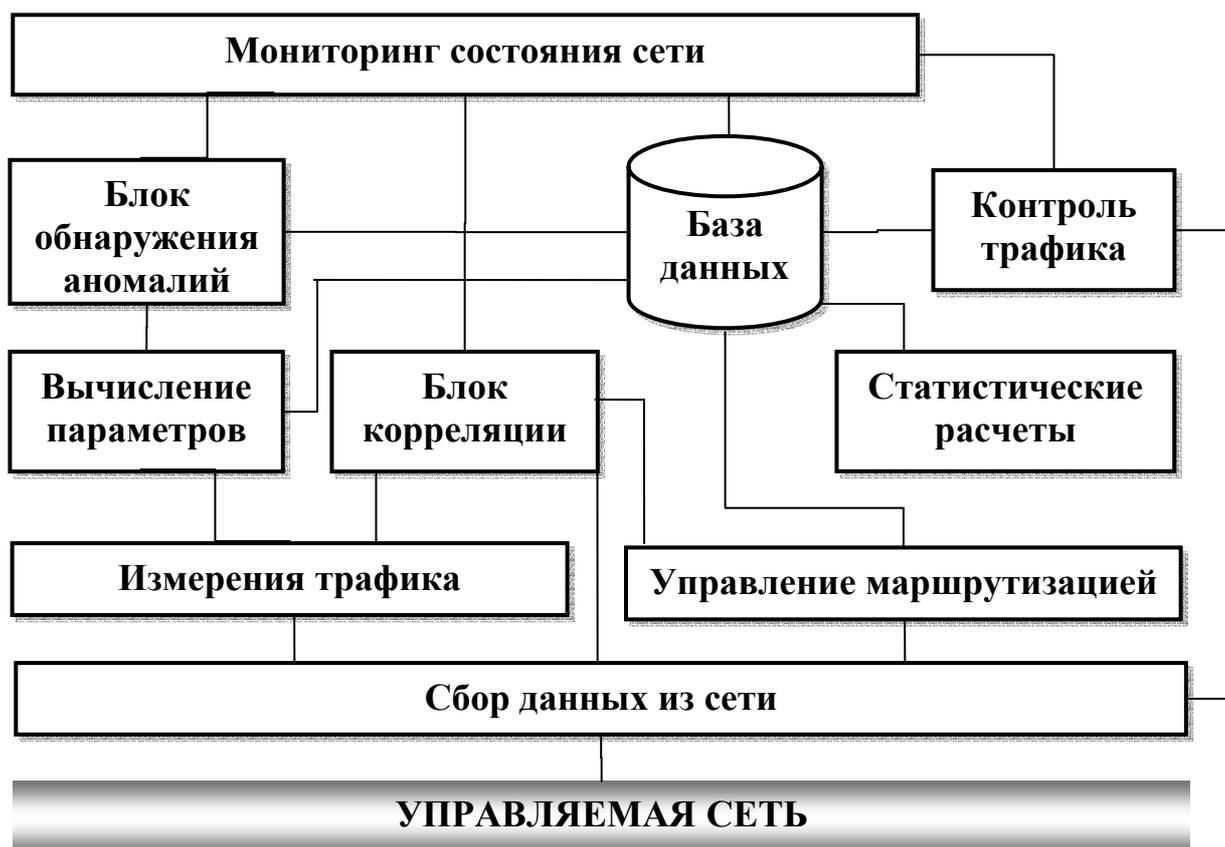


Рис. 7. Основная функциональная архитектура блока управления трафиком.

На основании сделанного анализа были решены задачи, соответствующие выполнению пунктов 1–4 общей методологии, описанной в предыдущем разделе. В частности, основная последовательность моментов принятия оперативных решений (масштаб времени) была определена как последовательность моментов появления вызовов. Глобальное «действие» на одном такте заключается в указании маршрута соединения для поступившего вызова, ли-

бо в отказе на соединение. Это действие сопряжено с рядом промежуточных решений, то есть имеет место декомпозиция процесса принятия решений, которая получается естественным образом и связана с технологией организации трафика в сети. На рис. 8 и 9 показана схема управления установлением соединений. Маршрут соединения отдельного вызова прокладывается постепенно, от узла возникновения вызова к узлу назначения. Локальное решение в каждом узле заключается в выборе пучка продолжения маршрута. Это решение принимается на основании маршрутной таблицы, соответствующей данному вызову, а также в зависимости от дополнительных управляющих воздействий, которые могут быть осуществлены в данный момент. В свою очередь параметры маршрутных таблиц и дополнительных управляющих воздействий регулируются на более высоком уровне. Таким образом, принятие решений в системе происходит распределенным многоуровневым образом в асинхронном режиме. Общее количество локальных подсистем, требующих последовательного оперативного изменения параметров (с учетом количества узлов в сети и типов вызовов) достигает порядка десятков тысяч.

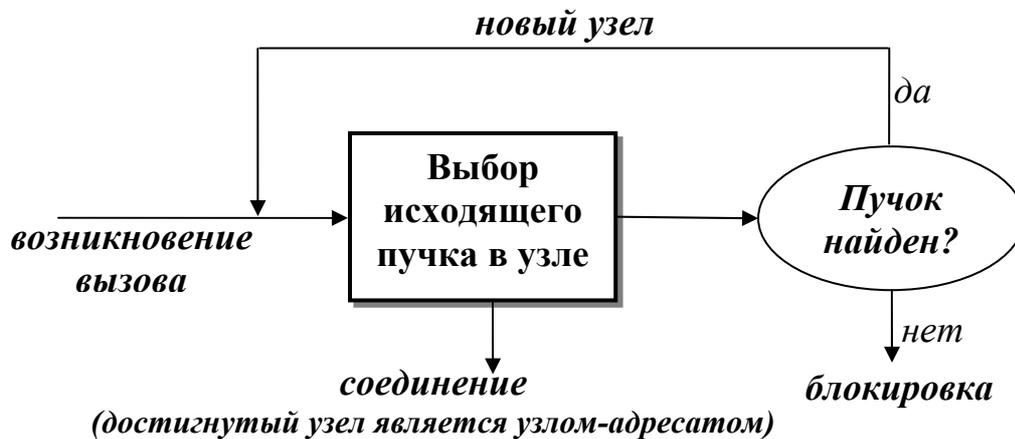


Рис. 8. Последовательное установление маршрута соединения.

В разделе 5.3 построена и проанализирована математическая модель управляемой сети с коммутацией каналов. Связанные с ней постановки задач обладают тем несомненным преимуществом, что допускают аналитическое решение. Модель дает возможность осуществить первоначальное исследование объекта, пренебрегая некоторыми реальными, но трудно формализуемыми факторами.

Модель базируется на предположениях об экспоненциальном характере распределений и названа *эрланговской*. Сеть представляется в виде множества узлов, соединенных между собой пучками каналов (линиями). Пучок каналов l , $1 \leq l \leq L$, имеет емкость n_l . На узлы поступают пуассоновские потоки вызовов с параметрами λ_m , $m = 1, \dots, M$. При поступлении вызова в сети принимается одно из двух решений: либо этот вызов блокируется (получает отказ), либо ему предоставляется маршрут для соединения, продолжительность занятия которого для вызовов m -го типа имеет экспоненциальное распределение с параметром μ_m . Для каждой тяготеющей пары определено множест-

во из K_m возможных маршрутов соединения. Состоянием сети в момент $t > 0$ является вектор $x(t) = \{x_{mk}(t), 1 \leq k \leq K_m, 1 \leq m \leq M\}$, компонента $x_{mk}(t)$ которого означает число вызовов m -го типа, которые в момент t имеют соединение по маршруту k . Множество возможных состояний сети имеет вид

$$X = \left\{ x = (x_{mk}) : \sum_{m=1}^M \sum_{k=1}^{K_m} a_{mk}^{(l)} x_{mk} \leq n_l, 1 \leq l \leq L; x_{mk} \geq 0, 1 \leq k \leq K_m, 1 \leq m \leq M \right\},$$

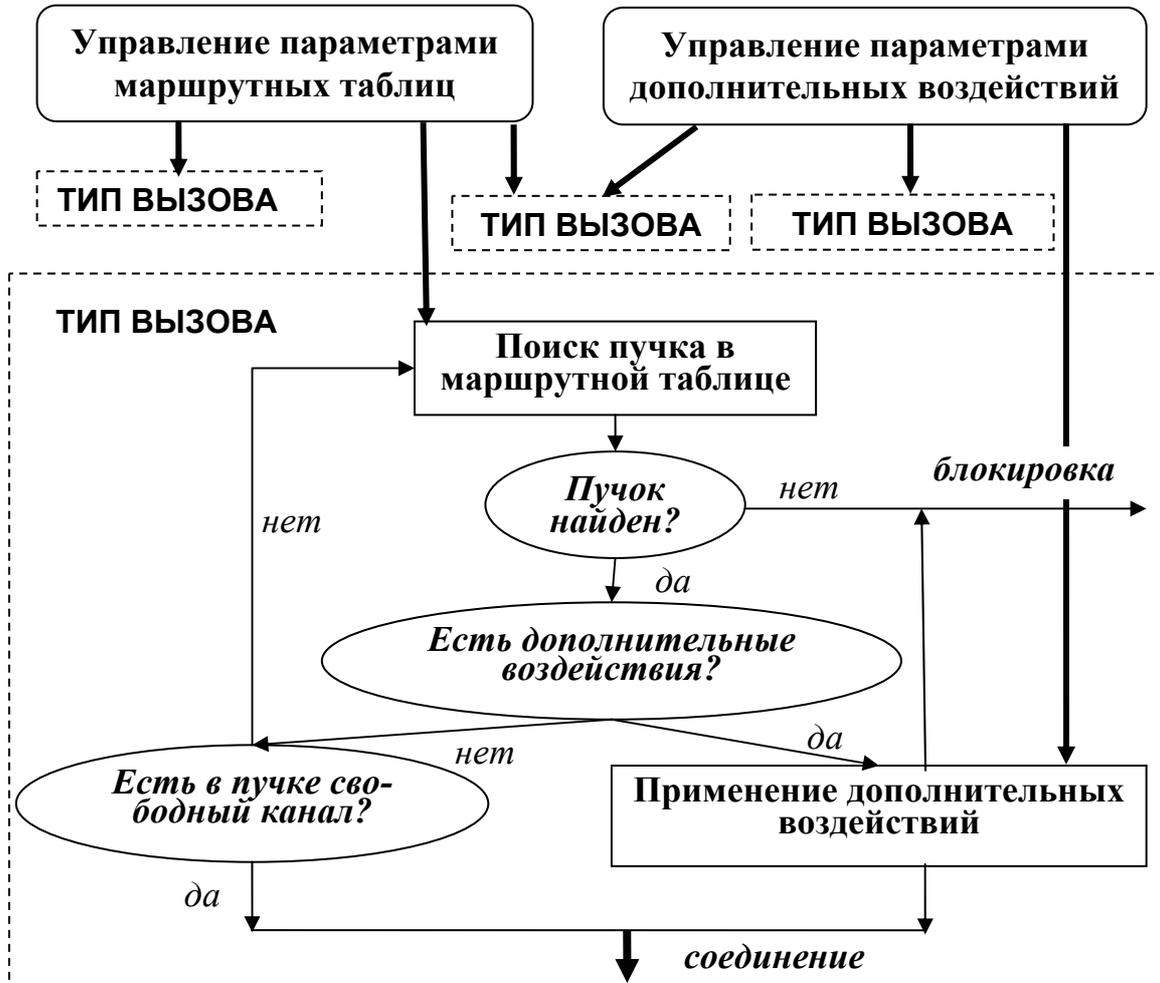


Рис. 9. Поиск пучка для продолжения соединения вызова данного типа.

где $a_{mk}^{(l)}$ равно 1, если линия l встречается в k -м маршруте соединения для m -го типа потока, и 0 – в ином случае. Соединение вызовов осуществляется с помощью набора векторов $u = \{u_{mk}, 1 \leq k \leq K_m, 1 \leq m \leq M\}$, где компонента u_{mk} означает вероятность выбора k -го маршрута для вызова m -го типа. Множество допустимых стратегий имеет вид

$$U_\delta = \left\{ u : \sum_{k=1}^{K_m} u_{mk} = 1, u_{mk} \geq \delta; 1 \leq m \leq M \right\},$$

где $\delta > 0$ – заданное число. Функция относительных предельных потерь для m -го типа потоков определена как $\Phi_u^{(m)} = \lambda_m^{-1} \lim_{t \rightarrow \infty} t^{-1} Z_u^{(m)}(t)$, где $Z_u^{(m)}(t)$ означает математическое ожидание суммарного числа блокировок для вызовов m -го типа за время от 0 до t по мере P_u , и может быть также выражена в виде $\Phi_u^{(m)} = \lambda_m^{-1} P_u^{(m)}$, где $P_u^{(m)}$ – предельная вероятность того, что вызов m -го типа получит отказ.

Теорема. *А) Если маршрутизация осуществляется с помощью стратегии $u \in U_\delta$, $\delta > 0$, то предельная вероятность нахождения процесса $x(t)$ в состоянии $x = \{x_{ij}^{(mk)}\}$ существует и равна*

$$\pi_x = \pi_{u,x} = \lim_{t \rightarrow \infty} P_u(x(t) = x) = \frac{(\rho u)^x}{Cx!},$$

$$\text{где } \tilde{N} = \sum_{x \in X} \frac{(\rho x)^x}{x!}, (\rho u)^x = \prod_{m=1}^M \prod_{k=1}^{K_m} (\rho_m u_{mk})^{x_{mk}}, x! = \prod_{m=1}^M \prod_{k=1}^{K_m} x_{mk}!, \rho_m = \frac{\lambda_m}{\mu_m}.$$

Б) Каждая точка локального минимума функции $\Phi(u) = \sum_{m=1}^M \Phi_u^{(m)}$ на множестве U_δ является точкой глобального минимума. Множество точек минимума функции $\Phi(u)$ связно.

В) Частные производные функции $\Phi(u)$ задаются формулой

$$\frac{\partial \Phi(u)}{\partial u_{lk}} = \lim_{t \rightarrow \infty} \sum_{m=1}^M \rho_m \text{cov}(x_m(t), x_{lk}(t)),$$

$$\text{где } x_m(t) = \sum_{k=1}^{K_m} x_{mk}(t), \text{cov}(x_m(t), x_{nk}(t)) = \sum_{x \in X} x_m x_{nk} \pi_x(t) - \sum_{x \in X} x_m \pi_x \sum_{x \in X} x_{nk} \pi_x(t).$$

В разделе 5.3 получен также сходный результат для распределенных стратегий маршрутизации, которые используют прием, известный в телефонии, как «откат». Для отыскания минимума функции потерь $\Phi(u)$ на множестве стратегий U_δ предложено использовать градиентную стратегию, для которой доказана **теорема**, устанавливающая ее оптимизационные свойства.

Разрабатываемая комплексная модель вторичной телефонной сети предназначена для включения в общий блок управления трафиком. В ее задачу, в частности, входит анализ ситуации и выработка рекомендаций по управляющим воздействиям в сети. Существенной особенностью является то, что модель рассчитана на работу в реальном времени. При этом информационный обмен между моделью и остальным блоком управления трафиком осуществляется каждые 15 минут. Иными словами, по истечении каждых 15 минут комплексная модель получает из основного блока новую информацию о сети и должна передать в основной блок подготовленные данные анализа и рекомендации по управлению.

Одним из существенных звеньев комплексной модели является компьютерная модель, имитирующая процесс поступления и обслуживания вызовов. Имитационная модель является тем промежуточным объектом, на котором осуществляется работа алгоритмов анализа и оптимизации перед тем, как передать результаты их работы в основной блок. При этом, естественно, особое значение имеет точность, с которой имитационная модель воспроизводит фактическую картину трафика в сети.

Основным и единственным "входом" имитационной модели (если не считать структуру сети) является матрица тяготения. В настоящее время не предусмотрены и не проводятся прямые измерения интенсивностей входных потоков, поэтому необходимо их реконструировать по имеющимся измерениям. Эта реконструкция должна занимать сравнительно небольшое время, для того чтобы в отведенный пятнадцатиминутный интервал успеть «запустить» имитационную модель и дать возможность отработать основным алгоритмам анализа и оптимизации. Так возникает дополнительная (по отношению к основным «оптимизационным» проблемам) задача идентификации параметров входных потоков по результатам косвенных измерений. В разделе 5.4 изложен алгоритм, разработанный для восстановления интенсивностей входных потоков по результатам измерений. Формальная постановка задачи предполагает, что (1) отказы на пучках независимы; (2) потоки, поступающие на пучки, являются пуассоновскими. При этих предположениях выведены рекуррентные процедуры расчета вероятностей отказа и потоков на пучках при заданных интенсивностях входных потоков. Идентификация интенсивностей входных потоков осуществляется в результате выполнения итеративной процедуры, которая минимизирует функцию невязки между измеренными и рассчитанными значениями вероятностно-временных характеристик.

В разделе 5.5 изложена формальная имитационная модель сети с коммутацией каналов. В качестве инструмента моделирования выбран аппарат теории параллельных процессов¹⁶, к числу достоинств которого относятся лаконичность, удобная для программирования форма представления, возможность легко изменять, добавлять и удалять отдельные блоки, наличие математических средств для проверки логической правильности модели.

Были учтены следующие особенности, присущие реальной сети:

- непуассоновские входные потоки и произвольные распределения времен ожидания;
- повторные вызовы в случае блокировки;
- ненадежность каналов;
- произвольное управление соединением, в т. ч. возможность отката и перенаправления, постановка в очередь;
- значимая затрата времени на соединение;
- сбор информации о сети и пр.

¹⁶ Ч. Хоар. Взаимодействующие последовательные процессы. – М.: Мир, 1989.

На рис. 10 приведены некоторые объекты имитационной модели, описанные на языке параллельных процессов.

```

СЕТЬ = ||  $m \in M$ .ТРАФИК ||  $k \in K$ .КАНАЛ || СЧЕТЧИКИ || ВРЕМЯ
M : ТРАФИК =  $v := v + 1$ ; ПАУЗА $_0(\Delta)$ ; (v.ВЫЗОВ || ТРАФИК)
X : ВЫЗОВ =  $v\_сеть \rightarrow$  УЗЕЛ
X : СОСТОЯНИЕ_ВЫЗОВА =
    ( $v\_сеть \rightarrow \mu := \langle I \rangle$ ;  $\kappa := 0$  |
      $v\_узел \rightarrow \sigma := \kappa.E$ ;  $\mu := \langle \sigma \rangle^{\wedge} \mu$  |
      $откат \rightarrow \sigma := \kappa.V$ ;  $\mu := (\mu')'$ ;  $\kappa := (\mu')_0$ ); СОСТОЯНИЕ_ВЫЗОВА
X : УЗЕЛ =  $v\_узел \rightarrow$ 
    ( $v\_канал \rightarrow$  УЗЕЛ | блокировка  $\rightarrow$ 
     ПОВТОРНЫЙ_ВЫЗОВ | откат  $\rightarrow$  ОТКАТ | соединение  $\rightarrow$ 
     ЗАНЯТИЕ)
X : ОЖИДАНИЕ =  $v\_узел \rightarrow$  (if  $\sigma.c \geq \sigma.C$  then (такт  $\rightarrow$  ОЖИДАНИЕ)
                               else ПОПЫТКА_СОЕДИНЕНИЯ)
X : ОБРАБОТКА_ВЫЗОВА =  $\rho := R(\sigma)$ ; case  $\rho$  of
    1:  $v\_канал$ 
    2: блокировка
    3: откат
    4: соединение
    end; КОНЕЦ
X : ПОВТОРНЫЙ_ВЫЗОВ = ( $m, m.v$ ).ВЫЗОВ || КОНЕЦ
X : ОТКАТ = if  $\sigma = I$  then ПОВТОРНЫЙ_ВЫЗОВ else ( $\mu := (\mu')'$ ;  $\sigma := \kappa.V$ ;
УЗЕЛ)
X : ЗАНЯТИЕ = ПАУЗА $_0(T)$ ;  $из\_сети \rightarrow$  КОНЕЦ
КАНАЛ( $k$ ) = ( $\{x.v\_канал, x \in X, k = x.\kappa\} \rightarrow$  КАНАЛ $_x(k)$ ) || ОТКАЗ_КАНАЛА( $k$ )
КАНАЛ $_x(k)$  = ( $\{x.из\_канала, k = x.\kappa\} \rightarrow$  КАНАЛ( $k$ )) || ОТКАЗ_КАНАЛА( $k$ )
ОТКАЗ_КАНАЛА( $k$ ) = сбой( $k$ )  $\rightarrow$  исправен( $k$ )  $\rightarrow$  КАНАЛ( $k$ )
X  $\cup$  M : ПАУЗА $_k(t)$  = такт  $\rightarrow$  (if  $k < t$  then ПАУЗА $_{k+1}(t)$  else КОНЕЦ)
ВРЕМЯ $_t$  = такт  $\rightarrow$  ВРЕМЯ $_{t+1}$ 

```

Рис 10. Фрагмент имитационной модели на языке параллельных процессов.

Раздел 5.6 посвящен детальному обсуждению механизмов принятия оперативных решений в сети и их параметризации, то есть представлению в том виде, который позволяет применить разработанные адаптивные алгоритмы.

Под оперативным решением понимается, прежде всего, выбор маршрута соединения вызовов, который осуществляется с помощью маршрутных таблиц. При существующем положении маршрутные таблицы являются объектами, которые составляются на этапе проектирования, а их непосредственное изменение в процессе эксплуатации является технически трудно осуществимой процедурой. Коррекция маршрутизации, необходимость в которой возникает, например, в случае возникновения нештатной ситуации на сети, осуществляется с помощью дополнительных управляющих воздействий. В то же время анализ этих воздействий показывает, что применение большинства из них фактически означает временное изменение маршрутной таблицы. По-

этому коррекцию маршрутизации удобно трактовать, прежде всего, как коррекцию маршрутных таблиц, возможность которой была включена в модель.

Разработано понятие *обобщенной маршрутной таблицы*, описание которой сводится к набору стохастических векторов $\mathbf{p} = \{p_0, p_1, \dots, p_K\}$. Каждый вектор p_i представляет собой вероятностное распределение на некотором конечном множестве и участвует определенным образом в «прокладывании» маршрута поступившего вызова (разным типам вызовов в различных узлах соответствуют разные обобщенные маршрутные таблицы). Фактическое «решение» – маршрут – появляется в результате выбора серии "вспомогательных" действий, выбираемых независимо друг от друга с помощью компонент набора \mathbf{p} . Такая ситуация укладывается в разработанную схему распределенного принятия решений при неполном наблюдении. Для оптимизации параметров обобщенных маршрутных таблиц используются адаптивные стратегии, построенные на основании полученных теоретических результатов.

Коррекция маршрутных таблиц является инструментом практически достаточным для того, чтобы реализовать функции регулирования трафика в случае возникновения внештатных ситуаций. Это положение подтверждается на многочисленных вычислительных экспериментах с компьютерной имитационной моделью. Однако поскольку на существующих станциях реакция на изменения в сети технически допустима в виде дополнительных управляющих воздействий, то при разработке комплексной модели решалась задача оптимизации этих воздействий.

Управляющими воздействиями в сети являются: блокировка кода, прерывание вызовов, пропуск, отмена направления, перенаправление каналов, временное перенаправление потоков по обходному пути, выборочное резервирование каналов. Общая схема оптимизации имеет тот же вид, что и для обобщенных маршрутных таблиц, и сводится к коррекции одного или нескольких стохастических векторов, параметризующих то или иное дополнительное управляющее воздействие.

Раздел 5.7 является интегрирующим в изложении комплексной модели управляемой цифровой сети с коммутацией каналов. В нем дается детальное описание структуры данных, входных параметров и основных блоков модели в том виде, который предшествует непосредственному написанию компьютерных кодов. Приведены принципиальные схемы имитации работы сети в двух режимах: без коррекции и с коррекцией маршрутных таблиц и управляющих воздействий. Подробно изложены наиболее важные блок-схемы и процедуры, имеющие принципиальное значение для качества и быстродействия программной системы, в том числе: создание и обработка вызовов, установление соединения, работа блока коррекции (рис. 11).

В **Приложении** содержится описание варианта программной системы, который послужил прообразом промышленной версии, переданной заказчику (ОАО Ростелеком). Кроме того, приложенный вариант явился инструментом для экспериментального исследования разработанной методологии и технологии синтеза адаптивных стратегий. На рис. 12 изображены фрагменты интерфейса этой программы.

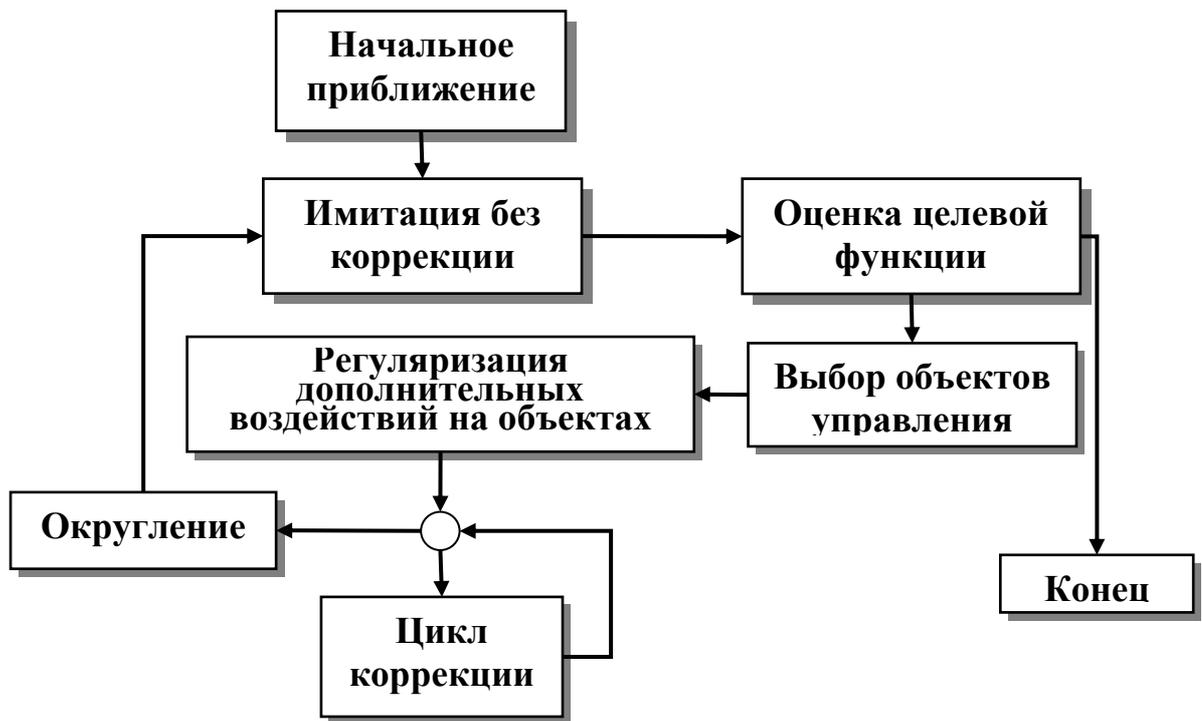


Рис. 11. Функциональная схема блока коррекции.

Компьютерная модель позволяет имитировать работу сети большого масштаба в двух режимах: с коррекцией и без коррекции маршрутизации и дополнительных управляющих воздействий.

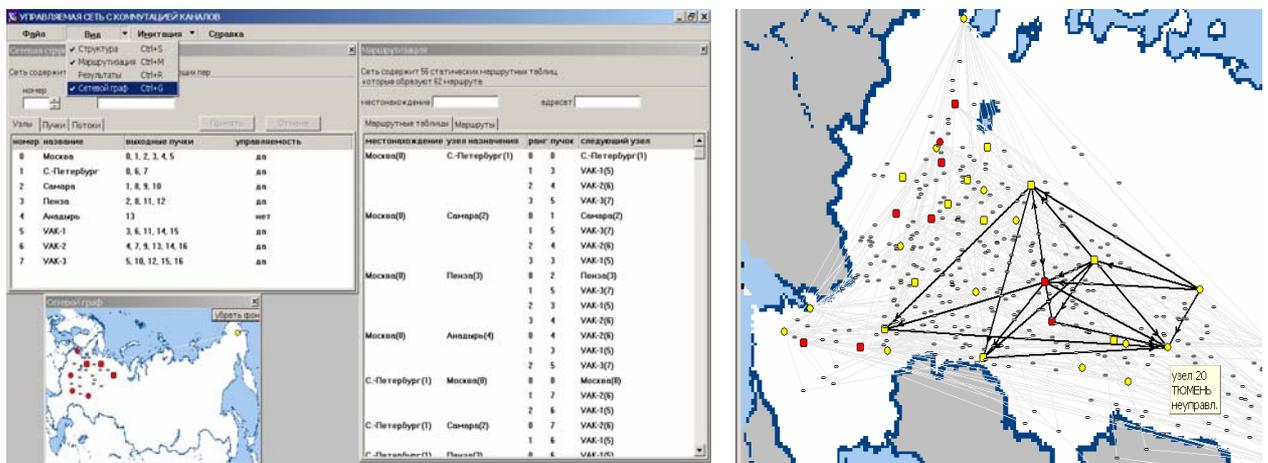


Рис. 12. Фрагменты интерфейса программной системы моделирования и оптимизации трафика междугородной телефонной сети

Представление о возможностях программы дает следующий пример. Входные данные примерно соответствуют параметрам телефонной сети России (на период 2000 г.). Сеть содержит 115 узлов, 2390 пучков и 6683 тяготеющие пары. Исходных маршрутных таблиц – 11140. Исходная маршрутизация также соответствует реальной (и, следовательно, была настроена на этапе проектирования и в ходе эксплуатации), поэтому можно считать ее близкой к оптимальной. В эксперименте проводилась коррекция 1110 управляемых маршрутных таблиц на управляемых станциях коммутации (без применения дополнительных управляющих воздействий), что соответствует на-

стройке более 10 000 параметров. Результаты эксперимента на ПК с тактовой частотой 2,4 ГГц:

- время выхода на стационарный режим – 2 мин;
- количество имитированных вызовов – около 10 млн.;
- относительное уменьшение вероятности отказов в сети после коррекции маршрутных таблиц – 6%.

Глава 6 посвящена описанию приложения разработанной методологии для синтеза информационной технологии моделирования и оптимизации системы распределенных вычислений.

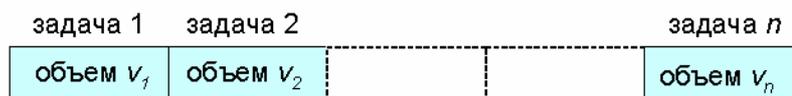
В основе проблематики лежит идея обеспечить коллективный доступ к вычислительным ресурсам для более эффективного их использования. Новые технологии в этой области выступают, главным (но отнюдь не единственным) образом, под общим названием «системы Грид». «Грид – это программно-аппаратная инфраструктура, которая обеспечивает надежный, устойчивый, повсеместный и недорогой доступ к высокопроизводительным компьютерным ресурсам»¹⁷.

Среди задач, возникающих в связи с проблематикой синтеза и эксплуатации систем распределенных вычислений и, в частности, технологии Грид, имеется универсальная проблема, присущая всем оттенкам концептуальных подходов и всем примерам практических воплощений подобных систем. Эта проблема заключается в поиске ответа на вопрос, как следует распоряжаться вычислительными ресурсами, то есть куда, в какие моменты и сколько следует посылать заданий, чтобы добиться эффективного использования ресурсов с точки зрения различных потребителей, а также с точки зрения различных показателей производительности ресурсов. Решение этой проблемы, по крайней мере, в части оперативного управления ресурсами, видится в применении адаптивных методов.

В **разделе 6.1** приведено описание модели оперативного управления распределением ресурсов на стадии непосредственного выполнения заданий, поступающих от потребителей. Основными объектами модели являются: заданное число распределенных вычислительных ресурсов, фиксированное число удаленных от ресурсов потребителей и потоки заданий, которые возникают у потребителей и которые состоят из случайного количества задач, упорядоченных с точки зрения очередности их выполнения (рис. 13). В модели учитывается сетевой аспект тем, что при передаче пакетов от потребителей к ресурсам и обратно происходят задержки, которые зависят от объема передаваемой информации и от «расстояний» между потребителями и ресурсами. Описаны механизмы возникновения заданий, отправки их на ресурсы и получения обратно, выполнения заданий на вычислительных мощностях.

¹⁷ Фостер Я. Что такое Грид? Три критерия. http://www.gridclub.ru/library/publication.2004-11-29.5830756248/publ_file.

Задание состоит из упорядоченного набора задач разного объема.



Задание выполняется путем посылки одного или более пакетов.

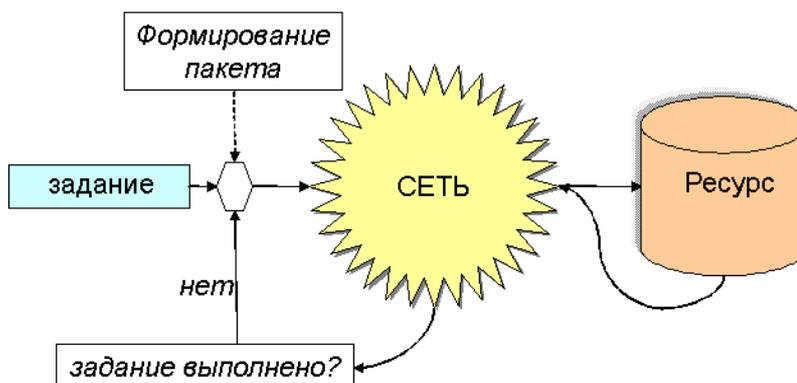


Рис. 13. Выполнение задания в модели сети Грид.

Раздел 6.2 посвящен постановке задачи и описанию адаптивных алгоритмов обработки информации для оперативного распределения ресурсов при неполной информации. Оперативные решения заключаются, прежде всего, в назначении ресурсов для задач, поступающих от разных пользователей. Кроме того, особенности, процесса обслуживания, в частности, возможность непредвиденного возвращения невыполненных задач, требуют оптимизировать объемы посылаемых пакетов задач. Оба типа решений представляются как выбор из конечного множества вариантов, и задача сводится к нахождению оптимальных значений распределений на этих множествах для различных пользователей и разных значений наблюдаемой части системы. Формально задача укладывается в схему частично наблюдаемой марковской последовательности и адаптивного распределенного принятия решений, которая была изучена в главе 3. В качестве основного одношагового показателя качества функционирования системы (одношагового дохода) выступает объем выполненных задач, возвращенных пользователю на данном такте. Соответствующим критерием качества решений является *производительность*, то есть отношение количества выполненных заданий к промежутку времени функционирования системы. Дополнительно рассмотрены одношаговые финансовые затраты, а также количество задач, потерянных (не обслуженных) на одном шаге. В этом случае получают целевые функции, обозначающие соответственно (предельные средние) затраты и (предельные средние) отказы. Рассмотрены два основных варианта принятия решений: централизованный, когда цель заключается в том, чтобы повысить среднюю производительность всей системы, и децентрализованный, когда каждый потребитель пытается увеличить собственную производительность. Адаптивные алгоритмы представляют собой рекуррентные последовательности, сконструированные в главе 3, в которых вероятностные распределения выбора тех или иных

управлений корректируются на основании накопленной статистики. Рассмотрены модификации адаптивных стратегий при различных предположениях о доступной наблюдению информации.

В разделе 6.3 приведены результаты вычислительных экспериментов с программной системой, в которой реализована модель распределенной вычислительной среды. Эта программа позволяет в широком диапазоне настраивать параметры модели, включая стратегии поведения участников системы, и проводить вычислительные эксперименты. Она является полезным инструментом для экспериментального исследования свойств системы, в частности, для сравнения свойств различных стратегий. Были изучены различные аспекты общих свойств адаптивных алгоритмов и поведения конкретной моделируемой системы. Рис. 14 позволяет увидеть типичную картину, когда адаптивная стратегия, после некоторого периода обучения в стационарной среде, выводит поведение системы на уровень, близкий к оптимальному. Надо подчеркнуть, что определение оптимальной стратегии иными, нежели чем адаптивными методами, возможно лишь в исключительных случаях. Причина в отсутствии аналитических методов и в большой размерности. Если добавить сюда принципиальную неполноту априорной информации и естественную неоднородность, присущую реальной системе, то использование адаптивного подхода выглядит практически неизбежным.

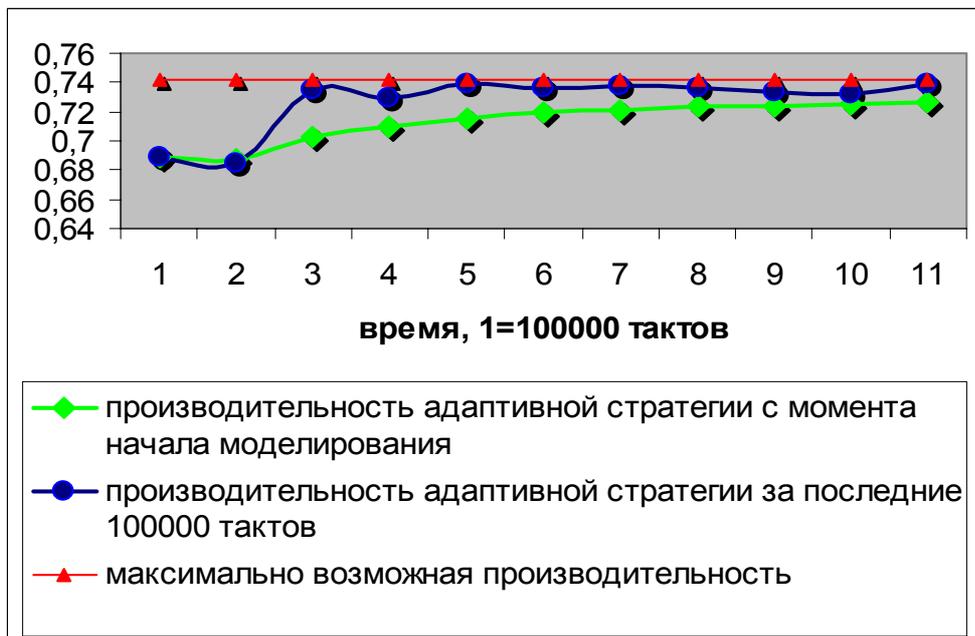


Рис. 14. Производительность адаптивной стратегии в сравнении с оптимальной стратегией (стационарная среда).

Ситуации с неоднородными во времени средами моделировались особо. На рис. 15 показан типичный пример, в котором на примере периодической среды сравнивается адаптивная стратегия и стратегия, которая представляется выгодной из соображений «здравого смысла». На рисунке различаются четыре зоны, которые соответствуют резко отличающимся свойствам системы. Эмпирическая стратегия (нижний график) оказывается более или менее эффективной на первых двух этапах, уступая немного адаптивной стратегии. На

следующих двух этапах преимущество адаптивной стратегии становится безоговорочным. На рисунке видны также области, на которых происходит приспособление адаптивной стратегии к резко изменившимся условиям.



Рис.15. Производительность адаптивной стратегии в сравнении с эмпирической стратегией (периодическая среда).

В результате проведенных экспериментов установлено, что адаптивные алгоритмы

- достаточно быстро находят оптимальную (неизвестную для них) стратегию поведения в тех случаях, когда оптимальная стратегия может быть определена априори;
- действуют результативнее, чем эвристические стратегии, в примерах, когда нет способа заранее и точно указать оптимальную стратегию;
- оказывают существенное положительное влияние на показатели качества, если их использовать для управления объемом посылаемых на обслуживание пакетов задач;
- являются восприимчивыми к выбору критерия качества;
- эффективны в случаях централизованного и децентрализованного (игрового) поведения участников.

РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ

В диссертации решена актуальная проблема анализа, синтеза и применения адаптивных алгоритмов оперативной обработки информации и принятия решений для широкого класса технических систем с неполным информационным описанием и неполным наблюдением. В том числе получены следующие результаты.

1. Разработаны теоретические основы адаптивной обработки информации и принятия решений в условиях неполного наблюдения.

1.1. Построена общая модель взаимодействия субъекта с частично наблюдаемым, априори неизвестным объектом в виде управляемой случайной по-

следовательности общего вида. Формализованы понятия объекта и стратегии (субъекта).

1.2. Дано определение класса регенерируемых объектов, который включает конечные частично наблюдаемые управляемые марковские цепи. Доказано, что для любого объекта из этого класса существует детерминированная однородная стратегия конечной глубины, которая является равномерно (относительно начального распределения) ϵ -оптимальной (по отношению к максимизации предельного среднего дохода).

1.3. Предложена конструкция адаптивной стратегии, основанной на идее «средних времен» и на переборе счетного множества вариантов (стратегия перебора). Стратегия перебора не использует априорную информацию об объекте.

1.4. Доказано, что стратегия перебора способна взаимодействовать с произвольным, априори неизвестным объектом не хуже, чем (неизвестная) наилучшая стратегия из произвольного заданного счетного множества стратегий.

1.5. Доказано, что стратегия перебора (на счетном множестве детерминированных однородных правил конечной глубины) обеспечивает максимально возможный предельный средний доход 1) для произвольного регенерируемого объекта, 2) для произвольной однородной конечной управляемой связной марковской цепи.

1.6. Доказано, что стратегия перебора обеспечивает для произвольного управляемого графа (с неизвестной структурой переходов и с неизвестным числом ненаблюдаемых вершин) предельный средний доход, равный тому максимально возможному, который можно получить в условиях полной информированности.

2. Разработаны методы адаптивной обработки информации для марковского процесса принятия решений.

2.1. Получен ряд свойств функции предельного среднего дохода в марковском процессе принятия решений со счетным множеством состояний.

2.2. Получены точные формулы градиента функции предельного среднего дохода для наблюдаемого и ненаблюдаемого распределенного марковского процесса принятия решений со счетным множеством состояний.

2.3. Предложена градиентная адаптивная стратегия для марковского процесса принятия решений с конечным множеством состояний и доказаны ее оптимизационные свойства.

2.4. Предложен способ построения асимптотически состоятельных оценок градиента целевой функции по результатам наблюдений в марковском процессе принятия решений.

3. Изучены возможности практического использования теории и указаны объекты приложений в областях инфотелекоммуникаций, производственных систем, искусственного интеллекта и других. Разработана методология практического применения информационных технологий моделирования и адаптивной обработки информации для широкого класса систем с неполным наблюдением. Разработана технология синтеза и реализации эффективных алгоритмов адаптивного принятия решений.

4. Получены результаты в области реализации разработанной теории и методологии для крупных прикладных проблем.

4.1. Решена задача оперативного управления трафиком в телекоммуникационной сети с коммутацией каналов. Разработана совокупность моделей и информационных технологий для оперативной адаптивной обработки информации в сети с коммутацией каналов.

4.1.1. Построена эрланговская модель управляемой сети с коммутацией каналов, в ее рамках поставлена и решена задача минимизации средних потерь в сети.

4.1.2. Разработаны модель и технология восстановления матрицы тяготения по результатам сетеметрии.

4.1.3. Построена имитационная модель трафика в сети связи с коммутацией каналов, написанная на языке параллельных процессов.

4.1.4. Дана трактовка модели управляемой сети с коммутацией каналов как управляемой марковской цепи с доходами. В этой трактовке введено и формализовано понятие обобщенных маршрутных таблиц, а также формализовано применяемое в практике телефонных сетей понятие «управляющих воздействий». Разработаны технологии адаптивной коррекции маршрутных таблиц и дополнительных решающих правил.

На базе полученных результатов создана программная система, реализованная в блоке управления трафиком междугородной телефонной сети.

4.2. Методология синтеза информационных адаптивных технологий и их применения в режиме off-line использованы в ходе проведения ряда ОКР для нахождения оптимальных системотехнических и программно-аппаратных решений при разработке АСУ и аппаратуры для инфотелекоммуникационных систем.

4.3. Создана информационная технология моделирования системы распределенных вычислений.

4.3.1. Построена информационная модель коллективного взаимодействия взаимно удаленных потребителей и распределенных вычислительных ресурсов.

4.3.2. Поставлена и решена задача адаптивного оперативного распределения ресурсов при неполной информации.

4.3.3. Экспериментально исследованы свойства адаптивных алгоритмов в модели системы распределенных вычислений.

Содержание диссертации изложено в следующих публикациях:

1. Коновалов М. Г. Об адаптивном управлении некоторыми классами марковских цепей // Доклады АН СССР. – Т. 233, № 5. – 1977. – С. 780–783.
2. Коновалов М. Г. Адаптивное управление неоднородными марковскими цепями // В сб. «Вопросы кибернетики. Адаптивные системы управления». – М.: Наука. – 1977.
3. Коновалов М. Г. Адаптивное управление периодическими процессами с независимыми значениями // Известия АН СССР. Техническая кибернетика, № 1. – 1979. – С. 138–144.
4. Гессель М., Срагович В. Г., Коновалов М. Г. Методы адаптивного управления конечными марковскими цепями и функционалами на них // В сб. «Вопросы кибернетики. Адаптивное управление». – М.: Наука. – 1979.
5. Коновалов М. Г. Об адаптивной маршрутизации в сети связи с коммутацией каналов // Известия АН СССР. Техническая кибернетика, № 3. – 1984. – С. 152–155.
6. Коновалов М. Г. Адаптивное управление конечными автоматами с ненаблюдаемыми состояниями // Доклады АН СССР. – Т. 291, № 1. – 1986. – С. 59–62.
7. Коновалов М. Г. Метод перебора в адаптивном управлении случайными процессами с дискретным временем. – М.: ВЦ АН СССР. – 1989. – 25 с.
8. Коновалов М. Г. Об управлении в сетях с коммутацией пакетов. – М.: ВЦ АН СССР. – 1989. – 40 с.
9. Коновалов М. Г. Опыт моделирования сети передачи данных на языке параллельных процессов // Математическое моделирование. – Т. 5, № 2. – 1993. – С. 82–93.
10. Коновалов М. Г. Адаптивная маршрутизация в сети связи с коммутацией каналов // В сб. «Всесоюзная конференция «Теория адаптивных систем и ее применение». – 1983.
11. Коновалов М. Г. Адаптивное управление конечными частично наблюдаемыми марковскими цепями // Четвертая международная вильнюсская конференция по теории вероятностей и математической статистике. – 1985.
12. Коновалов М. Г. Параллельные процессы и имитационное моделирование сетей связи // В сб.: XLVII научная сессия, посвященная Дню Радио. – М. 1992.
13. Коновалов М. Г. Адаптивный алгоритм маршрутизации для сетей передачи данных с коммутацией пакетов // В сб.: XLVII научная сессия, посвященная Дню Радио. – М. 1992.
14. Антонов С. В., Коновалов М. Г., Соколов И. А., Супрун А. П., Шоргин С. Я. Программные средства моделирования работы сетей с ретрансляцией кадров. Свидетельство об официальной регистрации программы для ЭВМ № 980363, РосАПО, 1998.
15. Антонов С. В., Коновалов М. Г., Соколов И. А., Шоргин С. Я. Математические методы и программные средства для анализа и проектирования сетей, использующих технологию ретрансляции кадров // Международная конференция по информационным сетям и системам ICINAS-98. Труды. – С.-Петербург: ЛОНИИС, ГУТ им. Бонч-Бруевича, РНТОРЭС им. Попова, 1998. – С.132–141.
16. Антонов С. В., Захаров В. Н., Коновалов М. Г., Шоргин С. Я. Комплексная модель цифровой телефонной сети и алгоритмы динамического управления // Системы и средства информатики. Вып. 11. – М.: Наука, 2001. – С. 68–77.

17. Коновалов М. Г. Управляемые марковские последовательности и оптимизация маршрутных таблиц в сетях связи с коммутацией каналов // В сб. «Системы и средства информатики», вып. 11. – М.: Наука, 2001. – С. 78–93.

18. Konovalov M. G. Management controls in telephone networks // 3 Московская международная конференция по исследованию операций (ORM2001). – Москва, 4–6 апреля 2001. – С. 57–58.

19. Шоргин С. Я., Соколов И. А., Антонов С. В., Захаров В. Н., Коновалов М. Г. Проблемы разработки математических методов и программных средств адаптивной оптимизации распределения потоков в многоуровневой сети коммутации каналов // *Обзорные прикладной и промышленной математики*. – Т. 9, Вып. 1. – 2002. – С. 252–253. (Третий Всероссийский Симпозиум по прикладной и промышленной математике (весенняя сессия)).

20. Соколов И. А., Антонов С. В., Захаров В. Н., Коновалов М. Г., Шоргин С. Я. Разработка математических методов оптимизации распределения потоков в многоуровневой сети коммутации каналов // *Обзорные прикладной и промышленной математики*. – Т. 9, Вып. 2. – 2002. – С. 452–453. (Третий Всероссийский Симпозиум по прикладной и промышленной математике (осенняя сессия)).

21. Коновалов М. Г. Восстановление матрицы тяготения по результатам измерений в сети // В сб. «Системы и средства информатики: Спец. выпуск. Математическое и алгоритмическое обеспечение информационно-телекоммуникационных систем». — М.: Изд-во ИПИ РАН, 2002. – С.100–111.

22. Соколов И. А., Антонов С. В., Захаров В. Н., Коновалов М. Г., Шоргин С. Я. Проблематика моделирования вторичных сетей телефонной связи // В сб. Труды конференции «Информационные технологии в науке, образовании, телекоммуникации, бизнесе». – Ялта–Гурзуф, 2002. – С. 125.

23. Коновалов М. Г. Экспериментальное сравнение некоторых алгоритмов маршрутизации в сетях с коммутацией каналов на примере сети Клоза // *Системы и средства информатики*. Вып. 13. – 2003. – С.106-121.

24. Антонов С. В., Захаров В. Н., Коновалов М. Г., Соколов И. А., Шоргин С. Я. Информационные технологии моделирования и динамического управления в многоуровневых сетях коммутации каналов // *Научно-технические технологии*, № 4. – 2003. – С.70-78.

25. Антонов С. В., Захаров В. Н., Коновалов М. Г. Имитационная модель в комплексной системе моделирования многоуровневой сети коммутации каналов. // Труды Международных научно-технических конференций «Интеллектуальные системы» (IEEE AIS'03) и «Интеллектуальные САПР» (СФВ-2003). – М.: Изд-во физико-математической литературы, 2003, Т.1. – С. 517-522.

26. Antonov S. V., Zakharov V. N., Konovalov M. G., Sokolov I. A., Shorgin S. Ya. Algorithms of routing and control actions optimization in multilevel networks of circuit switching // Seminar on Stability Problems for Stochastic Models. Universidad Pública de Navarra. Pamplona. 2003. P. 81.

27. Соколов И. А., Антонов С. В., Захаров В. Н., Коновалов М. Г., Печинкин А. В., Шоргин С. Я. Программные средства моделирования многоуровневой сети с коммутацией каналов и оптимизации распределений потоков. Свидетельство об официальной регистрации программы для ЭВМ № 2004611520, 21 июня 2004 г. (Федеральная служба по интеллектуальной собственности, патентам и товарным знакам.)

28. Соколов И. А., Антонов С. В., Захаров В. Н., Коновалов М. Г., Шоргин С. Я. Алгоритмическое обеспечение моделирования многоуровневой сети коммутации

каналов и оптимизации маршрутных таблиц // 8-я Международная конференция «Проблемы функционирования информационных сетей». Материалы конференции. МНПК «Связь-2004», Исык-Куль, 22-29 августа 2004 г. – С. 277–283.

29. Соколов И. А., Антонов С. В., Захаров В. Н., Коновалов М. Г., Печинкин А. В., Шоргин С. Я. Алгоритмы и программы моделирования и функционирования и процессов управления в телефонных сетях // Материалы международной научно-практической конференции «Информационные технологии и информационная безопасность в науке, технике и образовании ИНФОТЕХ-2004». Севастополь, 20–25 сентября 2004 г. С. 116–117.

30. Коновалов М. Г. Некоторые свойства функции предельного среднего дохода в задаче управления марковскими цепями // Вестник РУДН, серия Прикладная и компьютерная математика. – Т. 3, № 1. – 2004. – С. 61-77.

31. Коновалов М. Г. Об оценках градиента функции предельного среднего дохода в марковском процессе принятия решений // В сб. «Системы и средства информатики», вып. 14. – М.: Наука, 2004. – С. 68–85.

32. Коновалов М. Г. Градиентный подход к управлению конечными марковскими цепями и его применение // В сб.: «Проблемы и методы информатики. II Научная сессия Института проблем информатики Российской академии наук. Москва, 18-22 апреля 2005 г.». – М.: Наука, 2005. – С. 99–101.

33. Konovalov M., Shorgin S., Saverio S. Problems of GRID systems modeling. – Transactions of XXV International Seminar on Stability Problems for Stochastic Models. Maiori (Salerno), Italy, September 20–24, 2005. – P. 309.

34. Коновалов М. Г. Адаптивное управление процессом размещения заданий в модели коллективного использования распределенных вычислительных ресурсов. // В сб. Тезисы докладов 2-й международной конференции «Распределенные вычисления и Грид-технологии в науке и образовании». – Дубна, 26–30 июня 2006 г. – С. 76–77.

35. Антонов С. В., Душин Ю. А., Коновалов М. Г., Шоргин С. Я. Разработка математических моделей и методов распределения заданий в системе распределённых вычислений // "Системы и средства информатики". Выпуск 16. Москва. Наука. 2006 г. – С. 32–46.

36. Gaeta M., Konovalov M., Shorgin S. Development of mathematical models and methods of task distribution in distributed computing system // Reliability: Theory & Applications, Electronic Journal, ISSN 1932-2321. – Vol. 1, No. 4. – 2006. – <http://www.gnedenko-forum.org/Journal/index.htm>.

37. Коновалова И. Н., Коновалов М. Г. Модель и алгоритмы адаптивного управления для системы распределенных вычислительных ресурсов // Международный симпозиум «Информационные технологии и общество», 24 апреля – 01 мая 2007, Тель-Авив, Израиль. Материалы симпозиума. Москва, 2007. – С. 76–78.

38. Коновалов М. Г. Методы адаптивной обработки информации и их приложения. – М.: ИПИ РАН, 2007. – 212 с. – ISBN 978-5-902030-59-9.

39. Коновалов М. Г. Моделирование системы коллективного использования распределенных вычислительных ресурсов. Свидетельство об официальной регистрации программы для ЭВМ № 2008610082, 9 января 2008 г. (Федеральная служба по интеллектуальной собственности, патентам и товарным знакам.)

40. Коновалов М. Г., Чаплыгин В. В. Расчет матрицы тяготения по результатам измерений в сети с коммутацией каналов. Свидетельство об официальной регистрации программы для ЭВМ № 2008610083, 9 января 2008 г. (Федеральная служба по интеллектуальной собственности, патентам и товарным знакам.)