

СОДЕРЖАНИЕ

| | |
|---|-----|
| <i>Герасимов И. В.</i> Базовые принципы стратегических решений по становлению и развитию ФКТИ как системы предпринимательского типа | 3 |
| <i>Батищев Д. И., Власов С. Е., Балашов В. В.</i> Разработка алгоритмов распространения волны по макродискретам для однослойных структур | 17 |
| <i>Старостин Н. В.</i> Генетический алгоритм с комбинированным представлением решений для задач многокритериальной компоновки | 21 |
| <i>Батищев Д. И., Неймарк Е. А.</i> Решение задачи оптимизации нестационарной функции при помощи генетического алгоритма с использованием базы опыта | 29 |
| <i>Павлушин В. А., Михеев Ф. А., Марков М. В.</i> Генетические алгоритмы решения задач оптимизации частотных характеристик радиоэлектронных схем | 33 |
| <i>Иванов С. Г., Поздняков С. Н.</i> Компьютерное моделирование формирования математических понятий | 38 |
| <i>Васильев Н. В., Яшин А. И.</i> Использование алгебраического подхода для построения алгоритмов на графах | 43 |
| <i>Григорьев Ю. Д., Ле Динь Шон</i> О стратегиях управления рисками в задачах перестрахования | 47 |
| <i>Мотасем Абу Давас, Новакова Н. Е.</i> Интеллектуальная поддержка проектных решений с использованием технологий хранилищ данных и Data Mining | 52 |
| <i>Кутузов О. И., Татарникова Т. М.</i> Подход к Оптимизации структуры межсетевое устройства с привлечением генетических алгоритмов | 61 |
| <i>Романов А. Е.</i> Методика электронного макетирования электрических жгутов | 67 |
| <i>Комолова Н. В.</i> О методологии разработки стратегии управления предприятиями отрасли печати | 69 |
| <i>Хайсов Д. М.</i> Сетевое управление аппаратурой ВЧ-связи | 75 |
| <i>Куприянов М. С., Кочетков А. В.</i> Методы определения консистентного состояния в системах с самовосстановлением | 78 |
| <i>Ларистов А. И., Фахми Ш. С.</i> Пирамидально-рекурсивная обработка изображений | 85 |
| <i>Кюптер А. Б.</i> Введение в профессию | 88 |
| <i>Чирков А. Б.</i> Применение самоорганизующейся карты кохонена для решения задачи кластеризации | 90 |
| <i>Колодяжный А. Н.</i> Применение вейвлетов для выявления аномальных выбросов в траекторных измерениях | 95 |
| <i>Ветров А. Н., Котова Е. Е., Кузьмин Н. Н.</i> Адаптивная информационно-образовательная среда автоматизированного (дистанционного) обучения на основе параметрических когнитивных моделей | 101 |

Редакционная коллегия:

И. В. Герасимов
(председатель редколлегии)
Н. Б. Крупенко
(секретарь редколлегии)
А. И. Водяхо, Г. Д. Дмитриевич,
В. А. Егоров, М. С. Куприянов,
В. В. Цехановский,
М. Ю. Шестопалов,
В. В. Яновский

Редактор *Э. К. Долгатов*
Комп. верстка *А. П. Дробот*

Подписано в печать 10.10.06 г.
Формат 60x84 1/8.
Бумага офсетная.
Гарнитура «Таймс».
Печать: ризограф.
Печ. л. 13,75.
Тираж 100 экз. Заказ .

Издательство
СПбГЭТУ «ЛЭТИ»

197376, Санкт-Петербург,
ул. Проф. Попова, 5

© СПбГЭТУ «ЛЭТИ», 2006

БАЗОВЫЕ ПРИНЦИПЫ СТРАТЕГИЧЕСКИХ РЕШЕНИЙ ПО СТАНОВЛЕНИЮ И РАЗВИТИЮ ФКТИ КАК СИСТЕМЫ ПРЕДПРИНИМАТЕЛЬСКОГО ТИПА

Рассматривается идея «предпринимательского» университета, в ее рамках излагаются принципы, которые могут быть положены в основу концепции модернизации ФКТИ, отвечающей потребностям инновационной экономики.

Инновация, прагматическая модель специалиста, предпринимательская деятельность, образовательная услуга, человеческий капитал, практический интеллект, коллективная компетентность

Актуальность проблемы. Становление вуза как субъекта рыночных отношений связано в большой степени с вызовами внешней среды и с необходимостью поиска адекватных ответов на эти вызовы, сохраняя академические тенденции и ценности университетов и развивая их предпринимательскую культуру.

Эти процессы требуют определенных институциональных трансформаций традиционных университетов в направлении создания нового типа научно-образовательной структуры – инновационный университет.

Соответственно концепции научных, педагогических и технологических инноваций в совокупности с концепцией формирования предоставляемых образовательных услуг характеризуют общую стратегическую реакцию факультета на изменение внешних условий.

Цель статьи состоит в том, чтобы не просто познакомить читателя с идеями «предпринимательского» университета, но и определиться с ролью ФКТИ и возможными сценариями его модернизации, выработать согласованную позицию и в отношении принципов, которые следует положить в основу концепции модернизации ФКТИ, отвечающей потребностям инновационной экономики.

Проблемные ситуации в жизнедеятельности ФКТИ. ФКТИ осуществляет преимущественно фундаментальные и прикладные исследования, подготовку профессионалов повышенного творческого потенциала по научно-образовательному направлению «Информатика, управление и компьютерные технологии».

Информатизация как сложный социально-экономический процесс, разворачивающийся в современном обществе, определяет новый вид общественного разделения труда. Суть его – сбор, передача и переработка информации с целью обеспечения информационно-вычислительными услугами различных категорий пользователей. Отличительной особенностью этого вида труда по сравнению с другими является использование в качестве предмета труда принципиально нового ресурса общества – информации.

Информатика – научная база информатизации. Она способствует переходу от разрозненного использования методов и средств, изучающих структуру и общие свойства информации, а также закономерности ее создания, хранения, поиска, преобразования, передачи и использования в различных сферах человеческой деятельности, к целостной

информационной модели мира. Информатика интегрирует достижения теории познания, математической логики и вычислительной математики, кибернетики, лингвистики, семантики, системотехники, науки и техники о вычислительных машинах и организационной науки.

Интенсивное развитие информатики, наблюдаемое в последнее время, ее активное проникновение в смежные с ней и даже далекие области знания, ее обобщающий характер поставили высшую школу во всем мире перед сложнейшей проблемой организации преподавания этой науки. Два обстоятельства в первую очередь затрудняют ее решение. С одной стороны, необходимость обеспечить уровень образования, достаточный для целостного восприятия информационной картины мира, понимания путей использования методов и средств информатики в различных сферах деятельности, что является необходимыми атрибутами культурного человека. С другой стороны, узкая специализация профессиональной подготовки, которая позволила бы выпускнику вуза заниматься исследовательской, проектной, технологической, эксплуатационной деятельностью в условиях инновационной экономики.

Для ФКТИ речь идет о сохранении и развитии свойственных нашему вузу традиций подготовки *специалистов по разным направлениям как творческих личностей*. Отличительными признаками таких специалистов являются:

- способность к творческому и прагматичному мышлению, позволяющему адекватно оценивать складывающуюся жизненную и профессиональную ситуацию, к системному анализу данной ситуации и синтезу путей решения данной, ранее не встречавшейся проблемы;
- готовность к сотрудничеству, использованию инновационных идей, профессиональному развитию, управленческой и предпринимательской деятельности;
- владение полным набором средств общения (компьютер – как средство информационной коммуникативности, английский язык – как средство международной коммуникативности);
- противодействие набирающим силу негативным тенденциям деструктивного использования информации и информационных технологий.

Данный перечень содержит требования, которые напрямую не связаны с профессиональной компетентностью специалиста.

Отличительная особенность этих требований состоит в том, что они, в основном, характеризуют *сам процесс деятельности специалистов*. Последний непосредственно связан с функциональным структурированием их сознания (целенаправленным формированием информационных мозговых структур, на базе которых и осуществляется системное мышление).

Разрешение проблемных ситуаций (проблемные ситуации в жизнедеятельности ФКТИ представлены на рис. 1) означает *формирование новых механизмов жизнедеятельности, являющихся единством материальных, интеллектуальных и организационных условий новой деятельности, дополненным системой обеспечивающих эту деятельность ресурсов (и в первую очередь, информационных)*.

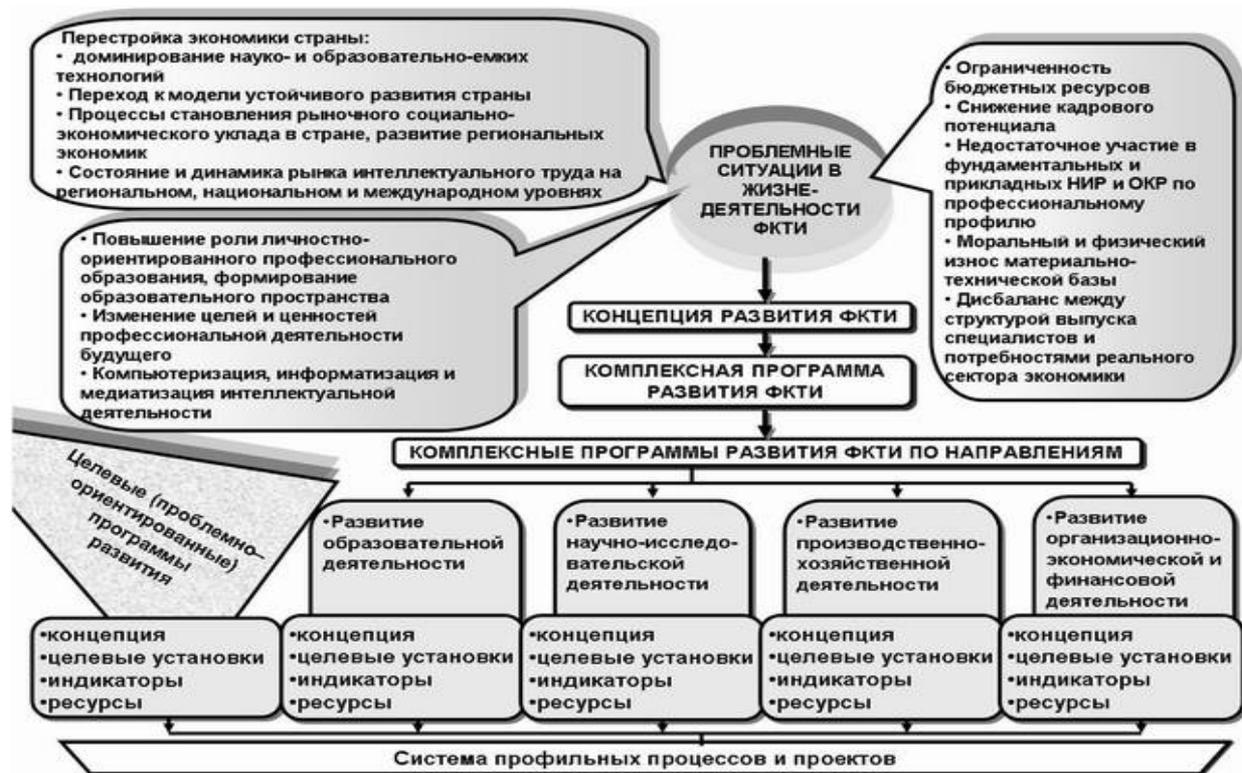


Рис. 1

Идея «предпринимательского» университета. Как показывает анализ проектов реорганизации университетов, модернизацию связывают в основном с идеями «предпринимательского» университета. Такой университет активно ищет инновации в своем бизнесе, вырабатывает существенные изменения в своем характере для того, чтобы занять более перспективные позиции в настоящем и будущем.

Прежде всего учитываются такие группы факторов, как рыночные отношения, изменение роли государства в развитии университетского образования; нарастающий темп изменений во всех сферах общественной жизни.

Сочетание данных факторов в их влиянии на развитие университета требует адекватной реакции на происходящие изменения как со стороны разработчиков образовательной политики, лидеров образовательных структур, так и студентов и преподавателей.

Рассмотрение происходящих изменений с двух позиций – системной и организационной – помогает оценить их характер.

| | | |
|----------------------|--|---|
| Уровень рассмотрения | Рыночные отношения, изменение роли государства | Нарастающий темп изменений (социальных, демографических, технологических) |
| Системный | Экономизация университета, стирание граней между предприятиями и университетом | Интеграция университетов и образовательных структур, конкуренция между университетами |
| Организационный | Создание адаптивных организационных структур | Коммерциализация технологий |

В условиях инновационной экономики кардинально изменяется окружающая внешняя среда, что ставит перед факультетом задачи не только адекватно отвечать на проблемные ситуации, но и искать *инновационные стратегические решения по становлению и развитию факультета как системы предпринимательского типа*, по изменениям в методах и организации менеджмента, в содержании и технологиях образования, взаимоотношениях с промышленностью, академическими институтами.

Можно утверждать о необходимости формирования *новой модели развития и пользования «человеческих ресурсов»*. В этой связи заслуживает внимания так называемая теория человеческого капитала, в соответствии с которой ресурсы, затраченные на образование, являются вложением в человеческий капитал.

Эта теория представляет большой интерес в плане понимания образовательных продуктов и услуг как товара (роль и место системы высшего профессионального образования (ВПО) в системе удовлетворения потребностей представлен на рис. 2). Она была разработана в рамках общепринятой теории предельной производительности для разрешения проблемного вопроса: «Почему разные люди получают разную заработную плату на определенном рынке труда, и почему структура оплаты труда на одних рынках существенно отличается от структуры заработной платы на других рынках?» и учитывала, кроме сугубо экономической рациональности, другие факторы. Основы этой теории были заложены еще классиками экономической теории и развиты современными экономистами.

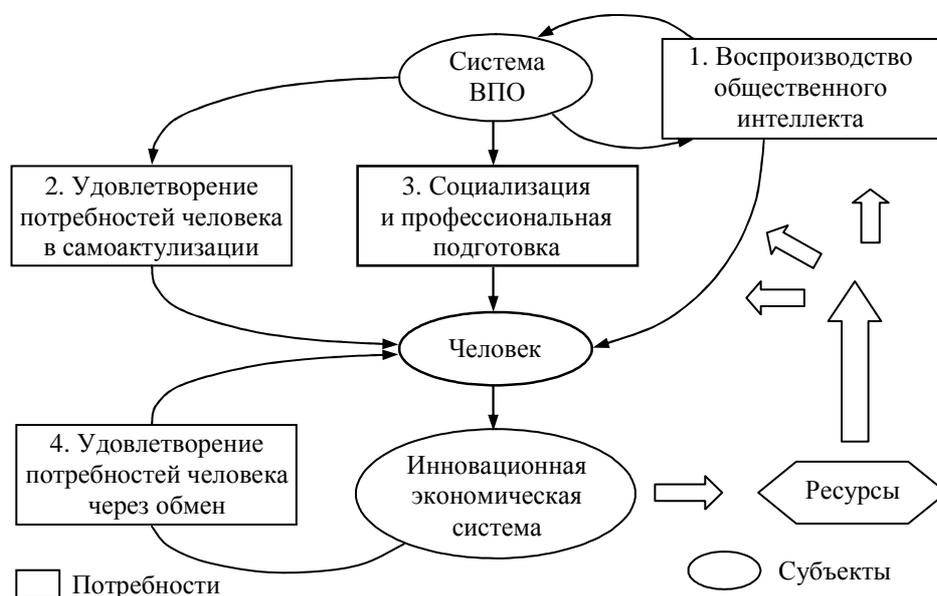


Рис. 2

В основе теории лежит положение, отвергающее тезис о том, что все люди созданы равными. Люди рождаются с определенными способностями или, по крайней мере, с необычными возможностями для их развития. В этом лежит причина различия в экономическом положении индивидуумов. Но способности, с которыми люди рождаются, – лишь полдела. **Обучение и воспитание – важны для многих специальностей не меньше, чем врожденная способность.** Данный факт дает возможность большинству людей активно влиять на свою судьбу. При этом затраты на образование рассматриваются как одна из форм инвестиций и должны быть достаточно велики, чтобы покрыть альтернативную стоимость обучения. Таким образом, систематическое образование является единственным способом вкладывать средства в человеческий капитал. Альтернативным систематическому может быть только практическое обучение в процессе работы (case studies). В целом увеличение человеческого капитала иллюстрирует рис. 3, где Ч – человеческий капитал (знания, умение, навыки)/цена рабочей силы; П – преподаватель; Н – наставник; ОУ – образовательное учреждение; t и Z – время и затраты на получение человеческого капитала на стадии систематического (с.о) и практического (п.о) обучения. Площадь круга условно отражает «размер человеческого капитала».

По сути дела, речь идет о переосмыслении поглощения системы ВПО в общественном восприятии и наращивании интеллектуального потенциала страны. Подчеркивая значимость государственного регулирования образовательных процессов, учитывая, что уровень государственных расходов является индикатором экономического развития общества, а задача воспроизводства общественного интеллекта остается одной из наиважнейших, мы все же исходим из положения, что учет рыночных категорий и процессов (индивидуальных и корпоративных потребностей, товара, спроса, предложения и др.) не только вполне применим к рассмотрению сферы образования, но и крайне продуктивен при формировании системы ее управления.

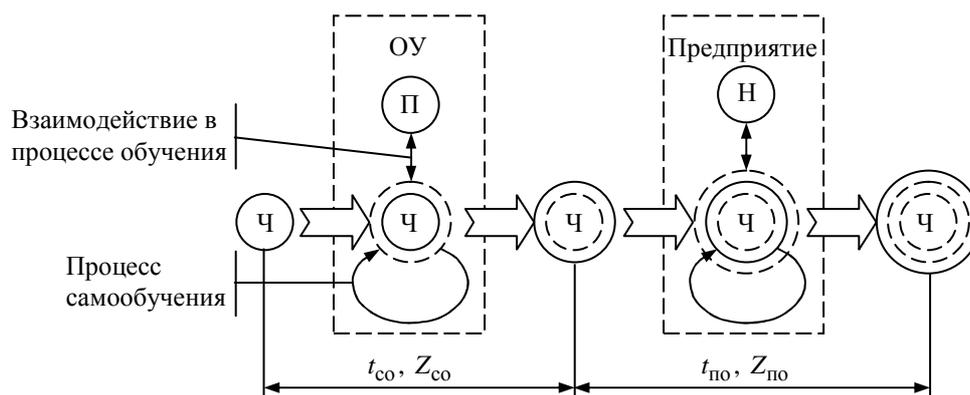


Рис. 3

В настоящее время предложение на отечественном рынке образовательных услуг в области новых информационных технологий обеспечивают: государственные образовательные учреждения высшего профессионального образования, негосударственные коммерческие высшие учебные заведения, а также коммерческие фирмы, предлагающие учебные курсы различной длительности и глубины по различным, пользующимся спросом на рынке, областям знаний. При этом важнейшим фактором, негативно влияющим на развитие государственных образовательных учреждений в последнее десятилетие, является острый недостаток госбюджетного финансирования. Это приводит к оттоку высококвалифицированных, опытных кадров и недостаточному пополнению кадровых ресурсов за счет молодых преподавателей – высшая школа стремительно «стареет». Данное обстоятельство неизбежно снижает ее конкурентоспособность на рынке образовательных услуг по новейшим направлениям и, таким образом, существенно сокращает возможности развития государственных вузов за счет привлечения дополнительных внебюджетных ресурсов. Вместе с тем, как показывает анализ, на отечественном рынке образовательных услуг имеется значительный и характеризующийся неуклонной тенденцией к росту платежеспособный спрос. При этом, как уже отмечалось, одним из самых динамичных и финансово емких сегментов данного рынка является подготовка специалистов в области новых информационных технологий, в частности, освоения новых программных продуктов. В настоящее время в этом сегменте рынка работает множество коммерческих фирм, предлагающих учебные курсы различной длительности и глубины по освоению конкретных программных продуктов. Такие фирмы и сертификационные центры ориентированы на коммерческие интересы конкретных производителей программного обеспечения и обучают работе с продуктами этих фирм-производителей, работая в жестких рамках их методик тренинга.

Основным противоречием современной системы образования следует считать противоречие между исключительно высокими темпами обновления знаний и, как следствие, спроса на рынке образовательных услуг, с одной стороны, и возможностями высшей школы в рамках существующих образовательных технологий и выделяемых госбюджетных ресурсов оперативно готовить и предоставлять широкий спектр востребованных обществом образовательных услуг высокого качества – другой.

В 2005 г. фонд «Центр стратегических разработок «Северо-Запад» осуществил комплексное исследование рынка образовательных услуг РФ на примере регионов, входящих в Северо-Западный федеральный округ (СЗФО). Установлен ряд тенденций, заслуживающих внимания в контексте обсуждаемого вопроса:

А. Растущая коммерциализация высшего образования. Высшее образование становится более коммерциализированным. Рост сектора платного образования и увеличение числа частных вузов идет в мире быстрыми темпами с 1980-х гг. В России платное образование появляется с середины 1990-х гг. и начинает развиваться быстрыми темпами.

Факты: если в 1995 г. число коммерческих студентов в российских вузах составляло 10 % от их общего числа, то уже к 2000 г. количество первокурсников-«бюджетников» и «платников» практически сравнялось. Из 1,4 млн. человек, поступивших в российские вузы в 2005 г., на бюджетных местах учатся 600 тыс. человек.

Б. Переход к «гибким» специальностям. В 1970–80-е гг. усложнение экономики привело к тому, что рынку труда стало требоваться все больше «гибких» специалистов. Они должны были выполнять работу, требующую владения разными отраслями знаний и умения приспособливаться к быстрой смене трудовых операций. Высшее образование отреагировало на это появлением большого числа новых специальностей, ростом междисциплинарных форм, а главное – переносом в процессе обучения студента акцента от *квалификаций к компетенциям*, что выражалось в способности выпускника принимать эффективные и оправданные решения в динамично меняющихся условиях. В России ориентация на «гибкие» специальности пришла на 1990-е гг. Специфика процесса заключалась в резкой «гуманитаризации» высшего образования, ускоренном импорте новых специальностей, особенно в сфере управления и услуг, росте числа новых специальностей на стыках старых дисциплин.

Факты: показатель увеличения спроса на «гибкие» специальности – российские работодатели при приеме на работу перестают ориентироваться на специальность кандидата, полученную в вузе. Опрос работодателей СЗФО показал, что только в половине компаний при проведении первичного собеседования с кандидатом просят показать диплом. На успеваемость кандидата в процессе обучения в вузе обращают внимание 68 % опрошенных компаний.

В. Открытие национального рынка образовательных услуг (интернационализация).

Интернационализация высшего профессионального образования началась в 1960-е гг.

Факты: в 2005 г. объем мирового рынка высшего образования достиг \$90 млрд. В 1996 г. в мире около 1,5 млн иностранных студентов, в 2005 г. – около 3 млн, а, по прогнозам, в 2010 г. – более 3,5 млн.

Доля России на мировом рынке образовательных услуг снижается. Причины: отсутствие конкурентоспособного продукта, слабая образовательная инфраструктура, отсутствие корпоративной позиции высшей школы на международном рынке образования.

Факты: 25-30 лет назад Россия делила с США 1–2-е место в мире по числу студентов-иностранцев, сейчас же находится на 8-м. Всего в России обучается 100 тыс. иностранных студентов (4...5 % от мирового спроса), и доходы от этого составляют примерно 1 % мирового оборота рынка образования.

Г. Появление новых форматов образования.

В России формирование этого рынка началось сравнительно недавно, но тем не менее можно говорить о некоторых новых услугах, влияющих на рынок.

Факты: опрос работодателей СЗФО показал, что в качестве методов развития персонала российские компании используют:

- 78 % – тренинговые программы;
- 52 % – семинары;
- 35 % – профильные курсы;
- 26 % – внутрифирменное обучение;
- 12 % – содействуют сотруднику в получении второго высшего образования;
- 9 % – содействие сотруднику в получении степени МВА;
- 6 % – иные формы повышения квалификации сотрудников.

Д. Смена образовательных технологий. Становится все более актуальной необходимость технологической революции в образовании, связанной с постепенным сокращением классно-урочного преподавания и ростом значения подготовки, основанной на использовании практического опыта (case-study), как способа подготовиться к трудовой деятельности.

Выявленные тенденции характеризуют причастность России к общемировому переходу сферы высшего образования к постиндустриальному типу. Хотя у нас этот переход начался всего 15 лет назад, все же он происходит слишком медленно. Он затянулся, и это дезориентирует людей в выборе вуза, в выборе специальности, в наборе квалификаций, предлагаемых на рынке образовательных услуг и не востребованных на рынке труда. Показатели, по которым идет отставание:

- развитие исследований в вузах;
- переход к «гибким» специальностям;
- развитие новых форм образования;
- внедрение новых образовательных технологий;
- рыночные позиции в мире и интернационализация российской высшей школы.

Резюме: возможные сценарные развилки развития системы ВПО России. Российское государство, определяясь в своем отношении к процессам, протекающим в национальной системе ВПО, должно будет пройти ряд сценарных развилок:

- Проводить кардинальную реформу ВПО или оставить систему эволюционировать, ограничившись общими реформами бюджетного финансирования вузов.
- Регулировать развитие ВПО в момент постиндустриального перехода (массовизация, платное образование, «гибкие» специальности и проч.) или сделать ставку на рыночное саморегулирование.
- Сосредоточиться в модернизации российской высшей школы на институциональных реформах либо стимулировать реализацию специально отобранных проектов (учебные программы, модули, округа, учебные заведения, технопарки, образовательные серверы, образовательная статистика, импорт технологий и проч.), на базе которых вырастить новую систему профподготовки.

- При реформировании перестроить всю существующую систему высшего профессионального образования или параллельно ей строить новую систему, не ломая сложившейся.

- Импортировать наиболее передовые образцы вузов или пытаться построить собственные.

- Выбрать в качестве основной точки концентрации ресурсов и усилий додипломное образование или финишные уровни подготовки, бизнес-образование, исследования, специализированное образование.

- Необходимо также определиться с социально-экономическими и территориальными приоритетами при развертывании новых блоков российской системы ВПО. Таких приоритетов сегодня видится 5:

- а) научно-технологический модуль;

- б) бизнес-образование;

- в) привлечение и натурализация мигрантов, удержание населения;

- г) достраивание конкурентоспособных кластеров российской экономики, причем сразу как интегрированных в мировую экономику (международные стандарты, формирование партнерской сети с ведущими корпорациями и проч.);

- д) региональное развитие.

Праксиологическая (деятельностная) модель выпускника. Традиционно университет объединяет под одной крышей много дисциплин, где разработка новых знаний и передача знаний интегрированы внутри академической организации. Несмотря на то что определенное согласование было и неизбежно останется на уровне факультетов или отделений, вполне возможно, что будущие задачи университета изменятся и приведут к иным требованиям качества, изменяющим связь в другом направлении, которое зарубежные исследователи называют «unbundling of the university» (развязка традиционных функций университета).

В инновационном университете должна обеспечиваться профильная подготовка инновационно-ориентированных специалистов в приоритетных областях науки, техники и технологий на основе единого процесса получения, распространения и применения новых знаний.

Как следует из изложенного ранее, инновационный университет – адаптивный к требованиям внешней среды, функционирующий и интенсивно развивающийся академический **комплекс коллективного предпринимательства**, действующий в конкурентной среде отечественных и зарубежных основных профильных рынков:

- подготовки специалистов интеллектуального труда;

- наукоемкой продукции и научного обслуживания;

- образовательных и консалтинговых услуг.

Деятельность инновационного университета ориентирована на удовлетворение потребительского спроса основных профильных рынков и получение максимальной прибыли, направляемой для повышения потенциала, возможностей сохранения своей академической сущности, своеобразия и индивидуальности, «моральной и интеллектуальной независимости от любой политической власти и экономической силы» (Великая хартия европейских университетов, 1988 г.).

Компетентный специалист в сфере инновационной экономики – профессионал, способный комплексно сочетать исследовательскую, проектную и предпринимательскую

деятельность, ориентированную на создание конкурентоспособной продукции. Это – создатель интеллектуальных ценностей, способный их реализовать и уверенный в успехе.

Компетенция – это способность инициативно и оперативно применять на практике знания, умения и навыки, полученные в ходе обучения; креативность и способность действовать эффективно в условиях неполноты информации и ограничений во времени; способность быстро переквалифицироваться в соответствии с требованиями рынка, способность выступить экспертом в конкретной предметной области; умение соотнести цель с применяемыми методами, выраженная способность выбрать те навыки, которые применимы к конкретной ситуации; способность увидеть проблему, готовность увидеть проблему, готовность сформулировать задачу, предложить ее решение и успешно ее решить*.

Заметим, что существует противоречие между формальной квалификацией молодого специалиста и его компетентностью. Нередко «заказчики» (государство, бизнес) недовольны уровнем компетентности «квалифицированного» выпускника. Требования современной фазы развития общества таковы, что *выпускник помимо основной квалификации должен обладать целым рядом иных «компетенций».*

В основу модели положен следующий тезис: *человек неизмеримо усилит свой интеллект, если сделает частью своей натуры способность планировать свои действия, вырабатывать общие правила и способ их применения к конкретной ситуации, организовывать эти правила в осознанную и выразимую структуру – одним словом, сделается программистом в широком смысле этого слова.*

Четыре важнейших аспекта модели заслуживают пристального внимания: когнитивный, коммуникативный, технологический и прикладной:

- *когнитивный аспект* дает представление о структуризации знаний в предметной области профессиональной деятельности и об окружающем мире, предполагает развитие новых математических и конструкторских методов для описания абстрактного мира науки, формирует на основе собственного опыта специалиста и получаемых им знаний представления о механизмах и законах восприятия и обработки информации человеком, технологическими и социальными системами;

- *коммуникативный аспект* отражает механизмы возникновения и развития интеллекта человека, связанные с коммуникативными и регулятивными процессами, определяет формы, используемые для передачи знаний от человека к человеку, специфицирует человеко-машинные интерфейсы для взаимодействия агентов в компьютерной сети, характеризует коммуникативный цикл; отражает синтез индивидуальных свойств и поведения участников совместной деятельности, исходя из заданной групповой динамики, определяемой отношениями кооперации и конкуренции, конфликта и сотрудничества, субординации и координации и пр.; определение состава участников и их ролей; принимаемая здесь базовая единица – участник в организации – предполагает расширение и смещение психологического базиса интеллекта от когнитивной психологии в психологию деятельности и социальную психологию и вообще – в социокультурную область;

* Квалификация – это степень подготовленности к какому-нибудь виду труда; профессия, специальность; набор знаний по комплексу дисциплин, заложенных в стандарт.

- *технологический аспект* характеризует новые технологические средства и методы, способные лучше традиционных решать задачи, связанные с интеллектуальными процессами;
- *прикладной аспект* предполагает рассмотрение конкретной работы по созданию конкурентоспособной продукции, коммерциализации научных и технологических разработок.

В дополнение полезно обратить внимание на следующее обстоятельство, связанное с особенностями предметной области деятельности факультета.

Информационные специальности приобретают массовый характер. Однако надо иметь в виду, что сейчас это, пожалуй, самая трудная из всех массовых профессий, причем, к сожалению, эта трудность не признана в должной мере.

Трудность заключается в том, что именно специалисты в области информатики и математики упираются в предел человеческого познания в виде алгоритмически неразрешимых проблем. Следует подчеркнуть, что трудности алгоритмизации зачастую не связаны с какими-либо запретами принципиального или теоретического характера, а обусловлены всего лишь ограниченными возможностями самого человека. Собственный стек программиста должен быть глубиной не в 5...6 позиций, как это обнаружили психологи у среднего человека, а той же, что и стек в его очередной задаче, подлежащей программированию, плюс еще 2...3 позиции. Трудность также и в том, что программист должен обладать способностью первоклассного математика к абстракции и логическому мышлению в сочетании с инженерным талантом сооружать любые алгоритмические конструкции из нуля и единицы. А кроме всего прочего, программист должен уметь слаженно работать в коллективе и многое другое. Информатика требует от человека несколько особого взгляда на мир, его потребности и эволюцию, особой моральной подготовленности к своему профессиональному долгу.

В этой связи особую важность приобретает внутреннее отношение человека к своему делу. Является ли компьютерная профессия специфической профессией? Это не риторический вопрос. Ответ на него имеет прямые организационные, юридические и образовательные последствия.

Принимая во внимание доводы, изложенные ранее, вполне логичны и такие вопросы: Как сочетать элитарность компьютерщика с массовостью подготовки к этому виду интеллектуальной деятельности? Как воспитывать программиста – через мировоззрение (университет) или путем профессиональных навыков (технический вуз)? Что такое индивидуальные способности в программировании, специфичны ли они и нужны ли? Эти вопросы являются частью общей проблемы профессиональной ориентации и подготовки по группе интеллектуально емких специальностей с обязательным учетом человеческого фактора.

Компетенции выпускника. *Специалист в сфере инновационной экономики – профессионал, способный продуктивно сочетать исследовательскую, проектную, производственную и предпринимательскую деятельность, направленную на создание социально-значимой продукции и оказание услуг на уровне мировых стандартов.* Речь идет о специалистах, характерной особенностью системы знаний которых является наличие прочного естественно-научного, математического и мировоззренческого фундамента, обеспечивающего деятельность в проблемных ситуациях, поиск множества вариантов решения задач, оптимальных решений удовлетворения потребностей потребителей.

Выпускники инновационного университета ориентированы не только на работу по специальности, но и способны успешно выступать в роли предпринимателей и создателей новых

рабочих мест. Это своего рода «синтетические» специалисты, хорошо знающие сферы бизнеса и профессиональную, имеющие специальную подготовку по трансферу технологий, включающую усвоение знаний и формирование способов системного проектирования, программирования роста и развития, стратегического менеджмента и маркетинга предпринимательской деятельности, овладение методами и средствами межкультурной коммуникации.

Превращение сложившейся системы инженерного образования в сферу освоения способов инновационной деятельности, коммуникативной, инженерной и предпринимательской культуры требует наполнения соответствующим толкованием таких компетенций, как

- социально-личностные,
- экономические и организационно-управленческие,
- общенаучные,
- общепрофессиональные (инвариантные к области профессиональной деятельности),
- специальные,
- правовые,
- креативные и т. д.

Эти компетенции должны обеспечить нашему выпускнику мобильность на рынке интеллектуального труда и подготовленность к продолжению профессионального образования.

Что же нового в изложенной трактовке «портрета» нашего выпускника?

Это 4 основных положения.

Первое. Основное внимание концентрируется на самых различных видах образовательной активности, когда мотивация к усвоению знания достигается путем выстраивания отношений между конкретным знанием и его применением в профессиональной либо предпринимательской деятельности.

Второе. Высокий уровень методологической культуры, превосходное творческое владение методами познания и деятельности, причем не только методами классического естествознания, ориентированными на поиск единственного решения, но и формулирование и широкое внедрение в образовательную культуру многокритериальной постановки и решения инновационных проблем, поиском множества вариантов решения задач, методов системного подхода к выбору оптимальных решений.

Третье. Владение полным набором средств общения (компьютер – как средство информационной коммуникативности, иностранный язык – как средство международной коммуникативности, здоровый образ жизни, спорт – как средство межличностной коммуникативности); готовность к сотрудничеству, использованию инновационных идей, управленческой и предпринимательской деятельности.

Четвертое. Ориентация при определении содержания компетентностей на общепринятые международные стандарты качества результатов образовательной деятельности.

Рассматривая новые и традиционные черты «портрета» нашего выпускника в аспекте их дополнительности и взаимоподдержки можно разработать функционально-целевую концепцию факультета, его обновленной структуры, сформулировать задачи развития. При этом речь должна идти не столько о выживании существующих структурных подразделений, сколько о становлении качественно новой системы инновационного профессионального образования.

Конечно, система такого образования должна быть ориентирована на перспективные тенденции, а практика – укоренена в академических и инженерных традициях СПбГЭТУ «ЛЭТИ». С этой точки зрения следует рассматривать и вопросы международной интеграции образовательной активности, вхождения в мировой рынок образовательных услуг в рамках Болонского процесса.

Базовые принципы стратегических решений по становлению и развитию факультета как системы предпринимательского типа.

Принципиально важным является понимание того, что ФКТИ предоставляет образование, призванное не столько удерживать, сколько способствовать как развитию его традиционных структурных единиц (кафедр), связанных между собой концептуально, содержательно, методологически и функционально, так и становлению новых нестандартных компонентов, обеспечивая их рациональную взаимообусловленность.

Главный принцип сотрудничества – вовлечение всех участников корпоративной деятельности в такие отношения, нарушение которых нанесет ощутимые потери (этические, интеллектуальные, финансовые и др.) каждому участнику.

Наше видение базовых принципов осуществляется с позиций интеграции имеющихся на факультете интеллектуального и информационного ресурсов, введения их в «предпринимательский оборот».

Принцип 1. Формирование инновационной корпоративной культуры факультета и внутренней конкурентной среды, обеспечивающих условия подготовки инновационно ориентированных специалистов.*

Конкуренты стремятся к совершенству и преимуществу достоинств в рамках системы взаимного сотрудничества. Функция конкуренции, как известно, состоит в том, чтобы поместить каждого участника социальной группы на то место, где он лучше всего может служить своей группе и каждому из ее участников. Это, по существу, общепризнанный метод отбора наиболее способных людей для исполнения должностного функционала. Там, где присутствует общественное сотрудничество, должен производиться определенный отбор. Конкуренции нет там, где распределение функций и задач между сотрудниками осуществляется исключительно административным путем.

Принцип 2. Диверсификация и индивидуализация структуры подготовки и оказания широкого спектра дополнительных образовательных услуг; преобладание высших ступеней и уровней образования.

Единство факультета проявляется в том, что образующие учебно-воспитательный процесс компоненты концептуально, содержательно, методически и функционально связаны в определенном логическом порядке. Благодаря этому *факультет оказывается той целостностью, которая со времени возникновения европейской университетской традиции образует основное звено системы высшего образования.*

До 2010 г. Россия полностью войдет в единое европейское пространство высшего образования. Наши студенты получают право участвовать в программах академической мобильности. Диверсификация системы образования связана с необходимостью придания ей

* Конкуренция (от лат. сталкиваться) – это экономическое состязание за достижение лучших результатов в области какой-либо деятельности.

большого разнообразия, расширения набора предлагаемых образовательных услуг, создания условий для построения индивидуальной образовательной траектории, наиболее полно отвечающей личностным образовательным запросам и личной деловой карьере.

Принцип 3. Ориентация на высокие наукоемкие образовательные технологии с использованием методов, основанных на изучении практики и работе в команде.

Достижения в области IT-технологий коренным образом изменили характер, способы и требования к подготовке современных специалистов. На передний план вышел управляющий аспект человеческой жизнедеятельности. Знания, с одной стороны, являются продуктом IT-технологий, а с другой – объектом их применения.

Возникла глобальная диверсифицированная информационная среда обитания, в которой определяющими стали процессы производства и потребления знаний. Осмысление этих изменений позволяет определить пути и способы преобразований в существующих системах профессиональной подготовки, охватывая весь образовательный период от начальной школы до структур повышения квалификации. Достигнутые результаты в области образования прежде всего технологизировали образовательный процесс и придали ему существенно информационный характер.

IT-технологии образования – это, прежде всего, методы достижения целей образования путем оптимизации процесса потребления информационного ресурса в ходе обучения и объективизации процесса контроля за ходом и результатами обучения.

Идея технологичности учебного процесса имеет большое значение для массового контингента студентов, так как одним из важнейших факторов функционирования высшей школы является ограниченность ресурсов.

Большим недостатком общеевропейской системы многие эксперты считают отсутствие серьезных навыков коллективной работы. С этим связано понятие «индивидуальной образовательной траектории». В рамках такой схемы организации учебного процесса даже понятие «студенческая группа» меняет свое значение. Это уже не коллектив, а совокупность индивидуумов, вместе слушающих один учебный курс.

Принцип 4. Ориентация на лучшие отечественные и зарубежные аналоги образовательно-профессиональных программ, международную аккредитацию реализуемых на факультете программ.

Существуют общепринятые мировые стандарты на образовательную деятельность. В мире давно сформировались структуры, которые сертифицируют специалиста, выдавая ему сертификат качества его знаний. Разработаны стандарты и для общеевропейского образовательного процесса. Полезно соотнести уровень подготовки наших студентов с принятым международным уровнем.

Принцип 5. Развитие инфраструктуры взаимодействия факультета с внешней средой, формирование стратегического партнерства с вузами, академическими и промышленными организациями, бизнесом.

Опасно, когда исчезают механизмы оценки нашей деятельности извне – можно потерять ориентир качества. Важна для факультета и всевозможная ресурсная поддержка со стороны предприятий и бизнеса, совместная реализация долгосрочных комплексных программ и проектов в научной, образовательной и инновационной сферах.

Создание открытой и многоукладной интеграционной инфраструктуры (инновационно ориентированного кластера) – системообразующая основа долгосрочного партнерства.

Принцип 6. Рыночные категории и процессы не только вполне применимы к рассмотрению сферы образования, но и продуктивны при формировании системы ее управления.

Успех администрирования в вузе основан на самом процессе управления, а не на амбициях отдельных людей, осуществляющих его. Администрация, профессорско-преподавательский состав и студенты должны рассматривать свой вуз как сообщество, члены которого приносят пользу друг другу.

Успех администрирования в вузе как социальном и культурном институте зависит от всех видов его деятельности и адекватности реакции на изменяющиеся воздействия со стороны внешней среды.

Акцентировка внимания только на внутренних проблемах может создать риск переоценки важности внутренних задач и видов деятельности факультета в ущерб его внешним задачам, успешное решение которых поддерживает и развивает традиции факультета, формирует его имидж, а также вполне определенный менталитет его выпускников.

Единство факультета обеспечивается представительными и административными органами управления, Уставом университета и договорными отношениями. Важнейшим фактором успешного развития факультета в современных условиях является гармоничное сочетание централизованного и децентрализованного управления. Усиливается роль механизмов самоорганизации и самоуправления в процессах возникновения и развития научно-образовательных коллективов, занимающихся инновационной деятельностью.

Заключение. Было бы нелепо надеяться, что возможно в столь сжатые сроки очертить в более или менее завершенном виде контуры не столь отдаленного будущего ФКТИ. Основные усилия были направлены на осмысление базовых принципов реорганизации нашей деятельности, характерных особенностей «портрета» выпускника ФКТИ исходя из потребностей инновационной экономики. То, что представлено в настоящих материалах, требует критического анализа и конструктивных предложений по совершенствованию. Конечный итог – базовые положения, которые легли бы в основу комплексной программы развития факультета. Последняя, представляя собой плановый документ, содержит объединенные глобальной целью комплексы целевых установок, соответствующих им индикаторов достижения целей, а также необходимую и достаточную для их реализации ранжированных по степени важности, упорядоченных по срокам исполнения, обеспеченных ресурсами систему проектов (мероприятий) учебно-методического, научно-методического, технологического, инновационного и организационного характера.

Основные наши усилия должны быть направлены на то, как сделать научно-образовательную среду факультета эффективной и устойчиво развивающейся, а значит, привлекательной сферой приложения талантов и инвестиций. Она должна стать источником непрерывного потока интеллектуальных продуктов, услуг, технологий, востребованных реальным сектором национальной экономики.

I. V. Gerasimov

BASIC PRINCIPLES OF STRATEGIC DECISIONS ON ORGANIZATION AND DEVELOPMENT OF FCSI AS SYSTEM OF BUSINESS TYPE

The idea of «business» university is considered, principles which can be put in a basis of the FCSI modernization concept adequate to needs of innovational economy are stated in its frameworks.

Innovation, praxiology model of the specialist, enterprise activity, educational service, the human capital, practical intelligence, collective competence

РАЗРАБОТКА АЛГОРИТМОВ РАСПРОСТРАНЕНИЯ ВОЛНЫ ПО МАКРОДИСКРЕТАМ ДЛЯ ОДНОСЛОЙНЫХ СТРУКТУР

Рассмотрены основные различия между классическим дискретом коммутационного поля и макродискретом. Предложены два алгоритма распространения волны по макродискретам собственной разработки. Один из них предпочтительнее использовать в начале трассировки, а второй – на этапе, когда часть трасс уже проложена и их рисунок имеет сложную конфигурацию.

Алгоритмы трассировки, волновой алгоритм, макродискреты, распространение волны по макродискретам, однослойная трассировка

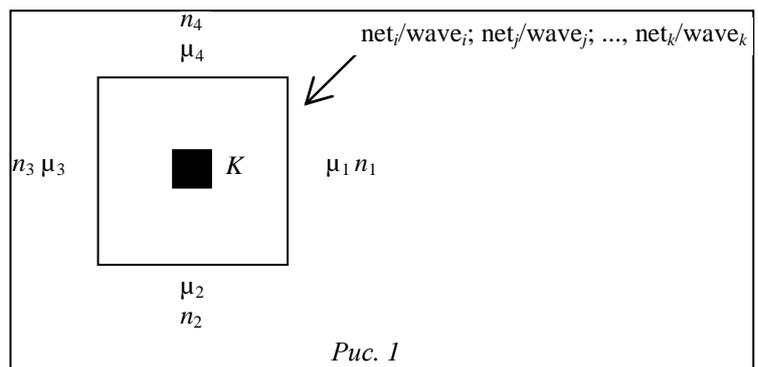
Большинство известных универсальных алгоритмов трассировки основаны на волновом алгоритме определения пути [1], [2]. При этом коммутационное поле разбивается на элементарные ячейки (дискреты), размеры которых определяются в соответствии с технологическими и топологическими ограничениями на допустимую плотность расположения проводников, выводов элементов и контактов между слоями металлизации.

Прямоугольный дискрет (рис. 1) характеризуется следующими параметрами:

- наличие в дискрете контакта K ;
- пропускная способность граней дискрета (μ_1, \dots, μ_4) , т. е. есть количество трасс, которые могут проходить через грани ячейки;
- количество трасс, проходящих через грань дискрета в текущий момент работы алгоритма трассировки (n_1, \dots, n_4) .

Каждая грань дискрета характеризуется «массивом цепей и волн», т. е. упорядоченным списком проходящих через нее цепей с соответствующими номерами фронтов волн ($net_i/wave_i; \dots, net_k/wave_k$).

В классическом определении сторона дискрета равна расстоянию между двумя соседними проводниками [1]. При делении плоскости монтажа на подобные дискреты число ячеек может быть очень большим, что требует большого объема памяти и довольно продолжительного времени на процесс распро-



странения волны. Распространение волны по укрупненным ячейкам (макродискретам) уменьшает требуемый объем памяти и время поиска трассы. Другим и более важным обстоятельством является то, что положение трассы в макродискрете фиксируется нежестко, и в дальнейшем при трассировке других проводников имеется некоторая свобода выбора [2]. Иногда метод трассировки по макродискретам называют гибкой трассировкой [3].

С точки зрения модели ячейки различие между классическим дискретом и макродискретом заключается в следующем:

1. Пропускная способность граней дискрета μ_i в классическом определении не может превышать 1:

$$0 \leq \mu_i \leq 1.$$

2. В макродискрете через грань может проходить более одной трассы, т. е. пропускная способность может быть больше 1:

$$0 \leq \mu_i \leq M.$$

3. Задачу трассировки по дискретам можно сформулировать как поиск такой геометрической формы соединений схемы, при которой загрузка граней всех ячеек не превышает допустимую:

$$n_i \leq \mu_i, \quad (1)$$

Т. е. количество трасс, проходящих через грани дискрета после завершения работы алгоритма, не должно превышать пропускной способности граней [4].

Алгоритмы волновой трассировки с использованием классических дискретов исследованы и подробно описаны в литературе. Алгоритмы распространения волны по макродискретам существенно отличаются от классических волновых алгоритмов. В литературе они практически не описаны. В данной статье приведены 2 алгоритма собственной разработки для распространения волны по макродискретам в однослойных структурах.

В классическом алгоритме фронт волны – характеристика дискрета, и каждому дискрету может быть поставлен в соответствие только один фронт волны. В случае использования макродискрета фронт волны (наряду с номером прокладываемой трассы) – характеристика ребер ячейки (рис. 1).

Так, например, у ребра *CG* (рис. 2) в момент работы алгоритма распространения волны для цепи 110 могут быть следующие значения «массива цепей и волн»: 110/2; 135; 110/2, где 135 – ранее проложенная цепь, а 110/2 – варианты прокладываемой цепи 110. На рис. 2 оба варианта цепи равнозначны. В дальнейшем один из вариантов может дойти до необходимого контакта 110, и тогда этот вариант фиксируется и становится проложенной цепью, а остальные варианты стираются.

Распространение волны в ребро возможно только при наличии в ребре минимум одного свободного канала, т. е.:

$$n_i - \mu_i \geq 0. \quad (2)$$

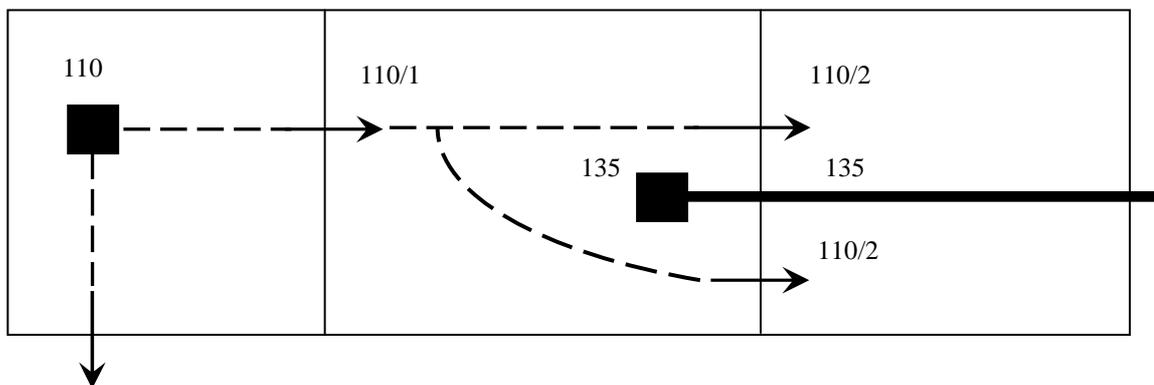


Рис. 2

Алгоритм № 1 распространения цепи S

1. Устанавливается текущий номер фронта волны $w_c = 0$.

2. В ячейке, в которой есть исходный контакт S, прокладываются все возможные варианты трасс из контакта во все грани ячейки (при выполнении условия (2)).

Для допустимых вариантов информация о трассах записывается в соответствующие грани ($net_i = S$, $wave_i = w_c$).

3. Текущий номер фронта волны w_c увеличивается на 1.

4. Просматривается все поле, и ищутся ячейки со значениями трассы в грани $net_i = S$, $wave_i = w_c$.

Если подобные ячейки найдены, то для каждой из них прокладываются все возможные варианты прохождения трасс в этой ячейке (при выполнении условия (2)). Для допустимых вариантов информация о трассах записывается в соответствующие грани ($net_i = S$, $wave_i = w_c$).

Если подобных ячеек не найдено, делается вывод о невозможности проведения трассы данным алгоритмом. Переход на шаг 7.

5. Проверка: дошла волна до необходимого контакта?

Если волна дошла, переход на шаг 6; иначе – переход на шаг 3.

6. Вариант трассы, дошедший до необходимого контакта, заносится в список проложенных цепей.

7. С поля стираются все варианты проложенных трасс.

Отличительными чертами данного алгоритма являются:

1) быстрота распространения волны по дискретному рабочему полю (ДРП);

2) в каждой ячейке может находиться только один фронт волны (аналогично классическому алгоритму);

3) во время этапа распространения волны варианты трасс проводятся через грань без точного учета свободных каналов, но необходимо наличие по крайней мере одного свободного канала, т. е. необходимо выполнение условия (2);

4) по ДРП одновременно может распространяться несколько вариантов трассы. Один из них может дойти до контакта. Остальные варианты стираются после завершения этапа распространения волны.

Таким образом, сделать вывод о том, можно ли проложить трассу этим алгоритмом или нет, можно распространив волну на все ДРП, при этом каждая макроячейка будет задействована максимум один раз.

Следует отметить, что на шагах 2 и 4 необходимым и достаточным условием для прокладывания вариантов трассы является условие (2), а выражение (1) не принимается в расчет. Фактически во время распространения волны пропускная способность граней может быть превышена. Впоследствии либо один из вариантов дойдет до нужного контакта-приемника (тогда этот вариант становится проложенной цепью, а все остальные варианты удаляются с ДРП), либо ни один из вариантов не дойдет до нужного контакта (в таком случае все варианты удаляются с поля и делается вывод о невозможности проведения трассы данным алгоритмом). В любом случае после завершения этапа распространения волны и удаления ненужных вариантов занятость каналов не будет превышать максимально возможную, т. е. условие (1) нарушено не будет.

Алгоритм №2 распространения цепи S

1. Устанавливается текущий номер фронта волны $w_c = 0$.

2. В ячейке, в которой есть исходный контакт S , строится список «доступных направлений», т. е. перечень всех возможных граней, в которые может быть проложена трасса (при выполнении условия (2)). Один из возможных вариантов прокладывается, и информация о трассе записывается в соответствующую грань ($net_i = S$, $wave_i = w_c$).

3. Текущий номер фронта волны w_c увеличивается на 1.

4. Просматривается все поле и ищется ячейка со значениями трассы в грани $net_i = S$, $wave_i = w_c$.

Если ячейка найдена, то для нее составляется список «доступных направлений», в которые можно проложить трассируемую цепь (при выполнении условия (2)). Один из вариантов прокладывается, и информация о трассе записывается в соответствующую грань ($net_i = S$, $wave_i = w_c$).

Если подобных ячеек не найдено, переход на шаг 6.

5. Проверка: дошла волна до необходимого контакта?

Если волна дошла, переход на шаг 7; иначе – переход на шаг 3.

6. В случае, когда распространение цепи дальше невозможно, необходимо вернуться на шаг назад, в предыдущую ячейку и попытаться провести трассу в ней иначе, используя массив «доступных направлений». Если массив «доступных направлений» в ячейке пуст, необходимо вернуться еще на шаг назад.

Если массив «доступных направлений» пуст, а алгоритм находится в исходной ячейке, делается вывод о невозможности проведения трассы данным алгоритмом. Переход на шаг 8.

7. Вариант трассы, дошедший до необходимого контакта, заносится в список проложенных цепей.

8. Все списки «доступных направлений» обнуляются.

Алгоритм № 2 имеет следующие особенности:

1) относительно медленное распространение волны по макродискретам коммутационного поля, так как приходится последовательно рассматривать все варианты прохождения волны в рамках всего ДРП. Причем если при использовании алгоритма № 1 за один шаг возможно распространение в несколько соседних ячеек нескольких вариантов цепи, то в алгоритме № 2 за один шаг возможно распространение только одного варианта только в одну соседнюю ячейку;

2) в одной ячейке могут находиться несколько участков одной и той же прокладываемой трассы с разными номерами фронта волны;

3) варианты трасс, прокладываемых через ячейку, сразу учитывают реальную пропускную способность граней;

4) по ДРП одновременно может распространяться только один вариант трассы.

В отличие от алгоритма № 1, в алгоритме № 2 постоянно учитывается реальная пропускная способность граней, т. е. если в грани один свободный канал (условие (2)), то через грань за все время работы алгоритма возможно проведение только одной цепи. Другими словами, в течение всего времени работы алгоритма нарушение условия (1) не допустимо. В этом ключевое отличие алгоритма № 2 от алгоритма № 1.

Еще одной особенностью алгоритма № 2 является то, что при его работе возможны «откаты назад» до последней «развилки».

Исходя из принципов работы алгоритмов, алгоритм № 1 предпочтительнее всего использовать в начале трассировки, а второй – на этапе, когда ДРП уже заполнено трассами и их рисунок имеет сложную конфигурацию.

СПИСОК ЛИТЕРАТУРЫ

1. Конструирование ЭВМ и систем: Учеб. для вузов. – 2-е изд., перераб. и доп. – М.: Высш. шк., 1989. – 312 с.
2. Абрайтис Л. Б. Автоматизация проектирования топологии цифровых интегральных микросхем – М.: Радио и связь, 1985. – 200 с.
3. Базилевич Р.П. Декомпозиционные и топологические методы автоматизированного конструирования электронных устройств. – Львов: Выща шк. Изд-во при Львов. ун-те, 1981. – 168 с.
4. Автоматизация проектирования БИС: В 6 кн.: Практ. пособие. Кн. 6. Казеннов Г. Г., Сердобинцев Е. В. Проектирование топологии матричных БИС/ Под ред. Г. Г. Казеннова. – М.: Высш. шк., 1990. – 112 с.

D. I. Batishev, S. Ye. Vlasov, V. V. Balashov

ALGORITHM DEVELOPMENT OF THE WAVE PROPAGATION IN MACRODISCRETES FOR SINGLE-LAYER STRUCTURES

The article reviews the main differences between the classic discrete of the connecting field and the macrodiscrete. There are proposed two in-house design algorithms of the wave propagation in macrodiscretets. It is better to use one of them at the beginning of the routing and the second one at the stage when a part of the routs have been routed and their connection layout has complex configuration.

Routing algorithms, wave algorithm, macrodiscretets, wave propagation in macrodiscretets, one-layer routing

УДК 621.372.001.24:681.3

Н. В. Старостин

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ С КОМБИНИРОВАННЫМ ПРЕДСТАВЛЕНИЕМ РЕШЕНИЙ ДЛЯ ЗАДАЧ МНОГОКРИТЕРИАЛЬНОЙ КОМПОНОВКИ

Рассматриваются проблемы компоновки, которые сводятся к многокритериальной декомпозиции взвешенных гиперграфов. Предлагается гибридная схема решения на базе генетического алгоритма. Обсуждаются особенности прямого и порядкового представления решений. Предлагается комбинированная символьная модель. Приводятся схемы оценки, улучшения и репродукции решений. Описывается механизм многокритериального отбора.

Компоновка радиоэлектронной аппаратуры, декомпозиция гиперграфов, генетический алгоритм

Задачи компоновки и математические модели. Проблема компоновки возникает при решении широкого спектра практических задач, таких, как унификация узлов радиоэлектронной аппаратуры (РЭА), распределение работ между исполнителями, размещение объектов и связей между ними, распределение частей программ между вычислительными системами и т. д. Многокритериальная особенность и сложный комбинаторный характер подобных задач приводят к немалым затруднениям в нахождении приемлемых решений¹.

¹ Речь о поиске точных решений в реальных задачах не идет из-за большой их сложности.

Классическими моделями для подобных задач являются графы, позволяющие сохранить информацию о схеме и строить формальные алгоритмы, удобные для реализации на вычислительных системах. Однако часто использование графов неудобно, поскольку предполагает введение избыточной информации и приводит к проблемам перехода от модели к моделируемому объекту. Ярким примером такого рода неудобств является проблема представления цепей в графовой модели при решении задачи компоновки элементов радиоэлектронной схемы [1]. Она решается либо заменой цепей полными подграфами, что вводит лишние ребра, либо заменой их произвольными деревьями, что случайным образом влияет на оптимальность общего решения задачи. При этом меняется содержательный смысл схемы, усложняются алгоритмы и возникает задача обратного перехода от графа к схеме из-за неоднозначности модели. В таких случаях гораздо удобней гиперграфовая модель [1]. Она точнее моделирует исходную схему и позволят избежать неоднозначности и избыточности. В частности, проблема представления цепей в описанной ранее задаче компоновки не существует в гиперграфовом представлении, поскольку цепи естественно моделируются гиперребрами.

Входными данными задач компоновки [1], как правило, являются: перечень и характеристики элементов, схема связей между элементами, условия совместимости элементов системы и декомпозиционные ограничения.

Решением задачи компоновки является распределение элементов по группам (классам, блокам, узлам). Решения, удовлетворяющие условиям совместимости и декомпозиционным ограничениям, называют допустимыми. Для оценки качества допустимых решений используют несколько критериев, направленных на минимизацию связей между группами элементов, группировку однотипных элементов, оптимизацию интегральных характеристик отдельных групп элементов и т. д.

Задачи декомпозиции гиперграфа. Формально задачи компоновки удобно сводить к задачам декомпозиции гиперграфа, в которых имеется взвешенный гиперграф $H(V, E, \varphi, \omega)$ порядка n , где $V = \{v_1, \dots, v_n\}$ – множество вершин; $E = \{e_i: e_i \subseteq V, i = 1..m\}$ – множество гиперребер; φ и ω – отображения, определяющие набор характеристик для каждой вершины и гиперребра соответственно. Требуется распределить множество вершин V гиперграфа $H(V, E)$ на k подмножеств V_1, \dots, V_k :

$$\begin{aligned} V_i &\neq \emptyset, \text{ для } i = \overline{1, k}; \\ V_i \cap V_j &= \emptyset \text{ для } \forall i \neq j, \text{ где } i, j = \overline{1, k}; \\ \bigcup_{i=1}^k V_i &= V. \end{aligned} \tag{1}$$

В зависимости от типа варьируемых переменных принято говорить, что задача декомпозиции имеет ту или иную степень свободы [2]. Задача с одной степенью свободы – число групп k и размер каждой группы $V_i = n_i, i = \overline{1, k}$ задаются перед решением задачи и в процессе выбора наилучшего разбиения остаются постоянными. Задача с двумя степенями свободы – число групп k задается перед решением задачи и в процессе выбора наилучшего k -разбиения остается постоянным, при этом порядки подгиперграфов разбиения могут варьироваться в заданных диапазонах. Задача с тремя степенями свободы – число групп k , размер каждой группы и сами разбиения являются переменными значениями.

Система требований (1) и условия, связанные со степенью свободы, могут дополняться другими ограничениями, например условиями совместимости элементов гиперграфа. Так, в задачах правильной вершинной раскраски графовых структур [3] таким ограничением является невозможность двух смежных вершин попасть в одну и ту же группу. Подобные требования, предъявляемые к разбиению (V_1, \dots, V_k) , определяют область поиска D .

Для оценки качества решений, как правило, предлагается некоторый набор критериев:

$$\begin{cases} \max_{x \in D} F_1(x), \\ \dots \\ \max_{x \in D} F_q(x). \end{cases} \quad (2)$$

Задачу, в которой присутствуют условие разбиения гиперграфа (1) и набор критериев (2) назовем многокритериальной задачей декомпозиции гиперграфа.

Схема решения на базе генетического алгоритма. Для решения рассматриваемого класса задач предлагается генетический алгоритм [1] – [5], дополненный специальными процедурами и операторами, в основе которых лежат эвристические подходы к построению и оптимизации структуры допустимых решений исходной задачи. В дальнейшем будем называть такой *алгоритм гибридным* [2].

В структуре гибридного алгоритма, по аналогии с генетическим, можно выделить 4 группы операторов: оператор формирования начальной совокупности решений, группа операторов по репродукции новых решений, операторы по оценке качества найденных решений и оператор отбора решений. Кроме того, гибридный алгоритм допускает включение в состав алгоритма различного рода специальных процедур и механизмов, главной целью которых является повышение эффективности поиска качественных решений.

Прямое вершинное представление решений. В задачах декомпозиции конечных гиперграфов имеется конечное множество допустимых решений D , что дает возможность решения $x \in D$ закодировать вектором $s \in S$ с фиксированным числом целочисленных компонент. Отображение, устанавливающее взаимно-однозначное соответствие между элементами множеств S и D , называют *символьной моделью*, или *представлением*.

На практике для рассматриваемого класса задач часто используют *прямое вершинное представление* [2]. Решение $x \in D$ представляется в виде строки длины n . Каждому символу в строке $(\square_1, \dots, \square_n)$ ставится в соответствие определенная вершина исходного графа: первый символ \square_1 соответствует вершине x_1 , \square_2 – вершине x_2 , ..., \square_n – вершине x_n . Значение символа будет указывать подграф, в который входит вершина.

К плюсам этого представления можно отнести алгоритмическую простоту, небольшую вычислительную сложность алгоритмов кодирования и декодирования. Кроме того, это представление позволяет применять различного рода эвристические подходы к построению или улучшению уже найденных решений в процессе генетического поиска [2]. Однако представленная модель обладает существенным недостатком. Классические генетические схемы репродукции решений в общем случае не исключают возможности появления строк, соответствующих недопустимым решениям задачи. Такого рода строки (кодировки) называют *некорректными*.

Решают проблему устранения из генетического поиска некорректных строк либо методами штрафов, либо используя специальные схемы коррекции и специфические операторы раз-

множения, эксплуатирующие дополнительную информацию о структуре получаемых решений. Но чем больше дополнительных ограничений в исследуемой задаче, тем сложнее построить удачные схемы генетического алгоритма на базе прямого вершинного представления.

Порядковое вершинное представление решений. Данное представление определяет порядок проверки вершин на возможность их включения в одну из групп. Некоторое решение задачи представляется в виде строки, состоящей из n символов, принадлежащих алфавиту $V_n = \{1, 2, \dots, n\}$. Каждый символ в строке обозначает номер вершины, а номер его позиции в строке определяет порядок проверки данной вершины. Такое представление подразумевает наличие некоторого конструктивного алгоритма, задача которого заключается в последовательном распределении вершин по подграфам. Одним из входных параметров такого алгоритма является строка, определяющая порядок проверки вершин. Результат работы алгоритма – некоторое допустимое решение задачи, т. е. такой конструктивный алгоритм работает в режиме декодера

$$(\square_1, \dots, \square_n) \square x,$$

и его предназначение заключается в построении допустимого решения по заданной строке-кодировке. Декодирование используется только для оценки получаемых строк-кодировок в процессе работы алгоритма.

Очевидно, что при порядковом представлении к строкам-кодировкам предъявляется требование, связанное с необходимостью перебора и проверки всех вершин в графе:

$$\square_i \square \{1, 2, \dots, n\}, i = 1..n;$$

$$\square_i \square \square_j \text{ для } i \square j \text{ при } i, j = 1..n.$$

Наличие библиотеки порядковых операторов репродукции новых решений упрощает реализацию общей схемы генетического алгоритма. Все, что необходимо – это наличие эффективного конструктивного метода решения задачи.

Главный недостаток данного представления заключается в том, что для ряда задач бывает крайне трудно, а иногда просто невозможно подобрать алгоритм кодирования $x \square (\square_1, \dots, \square_n)$, что жестко определяет конструктивные особенности схемы генетического алгоритма, не допускающей включения дополнительных эвристических процедур, направленных на улучшение уже найденных решений. Более того, при этом представлении не всегда удастся подобрать такой декодер, который гарантировал бы построение из множества всевозможных перестановок списка вершин любое решение x из множества допустимых решений D . Следовательно, не исключены ситуации, когда выбранный алгоритм декодирования позволит осуществлять поиск только в какой-то части из всего множества допустимых решений, поэтому к выбору конструктивного алгоритма нужно подходить очень серьезно. В качестве компромисса используют схемы, работающие с несколькими декодерами. Это приводит к некоторому усложнению общей схемы гибридного алгоритма, однако позволяет задействовать в исследовании новые области, что в некоторых случаях может повысить качество генетического поиска [3].

Комбинированное представление решений. Предлагается трехуровневое представление решений. Первый уровень – это уровень приложения генетических операторов. Здесь решение кодируется в виде перестановки π вершин исходного гиперграфа. Каждая

перестановка однозначно определяет результат работы декодера $x_{\pi} \in D$ – распределение вершин по группам, удовлетворяющее условиям разбиения.

Для класса задач декомпозиции графовых структур можно подобрать множество эвристических схем улучшения начальных решений по одному из q критериев оптимальности. Второй уровень несет в себе информацию о том, какая из улучшающих схем будет применена к начальному решению x_{π} . Примером кодировки может служить λ -ичная строка длины q . Каждой позиции соответствует оператор, улучшающий результат, полученный на первом уровне. Число в позиции определяет степень улучшения решения по заданному критерию.

Третий уровень – это уровень приложения эвристических операторов, определенных на предыдущем уровне и направленных на улучшение решения. Здесь решение кодируется прямым вершинным представлением. Этот уровень недостижим для генетических операторов кроссовера и мутации. На этом уровне работают эвристические схемы улучшения решения.

Структура трехуровневой символьной модели

| Уровень | Представление | Какие операторы и чем манипулируют |
|--|--|---|
| 1. Уровень работы генетических операторов | Порядковое вершинное представление в виде перестановки вершин гиперграфа | Порядковые операторы кроссовера и мутации формируют новые перестановки |
| 2. Уровень подбора эвристик | Код, описывающий последовательность применения улучшающих операторов | Многокритериальные стратегии подбора эвристик для нахождения области компромиссов |
| 3. Уровень работы эвристических операторов | Прямое вершинное представление в виде k -ичной строки длины n | Эвристические процедуры, направленные на улучшение решений по одному из критериев |

Таким образом, перестановка на первом и кодировка на втором уровнях однозначно задают схему получения допустимого решения задачи, закодированного в виде k -ичной строки на третьем уровне (см. таблицу).

Схема оценки кодировки. Оценка кодировки включает в себя 3 этапа. Первый заключается в восстановлении решения по перестановке, определенной на первом уровне символьной модели. Исходная перестановка задает порядок перебора вершин для некоторого конструктивного алгоритма, который последовательно формирует допустимое разбиение исходного гиперграфа.

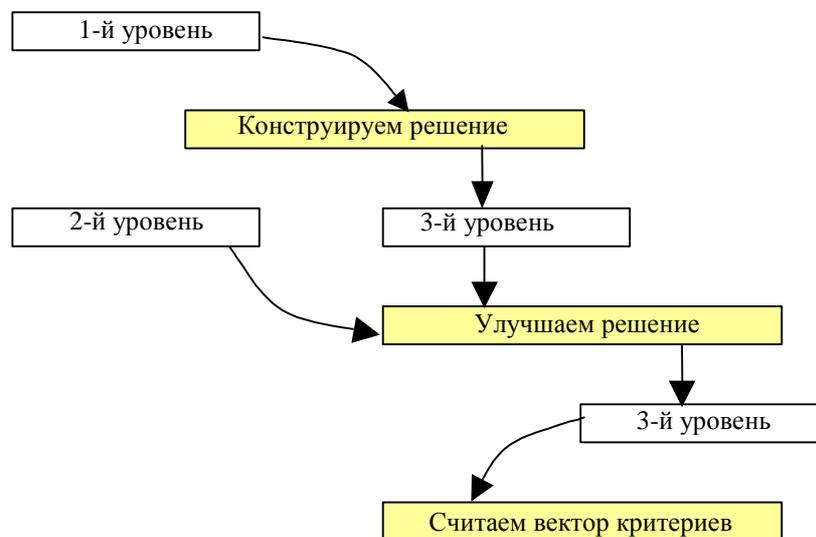


Рис. 1

Цель второго этапа – улучшить начальное разбиение по некоторому критерию посредством применения эвристик. То, какая из эвристик будет применена, задано в кодировке на втором уровне символьной модели.

Последний этап заключается в расчете вектора критериев построенного решения (рис. 1).

Процедуры улучшения решений. Предлагается оператор локальной адаптации, который осуществляет поиск некоторого локального оптимума по одному из частных критериев в окрестности относительного обмена [2] – всевозможных допустимых разбиений гиперграфа $H(V, E, \varphi, \omega)$, которые можно получить из текущего разбиения (V_1, V_2, \dots, V_k) с помощью обмена двух вершин разных подгиперграфов. Оператор может быть применен к заданному разбиению несколько раз путем повторения однократных обменов. Их число назовем *глубиной локальной адаптации*. Если разбиение после действия оператора адаптации невозможно улучшить с помощью однократных обменов, локальную адаптацию будем называть *полной локальной адаптацией*.

Как показывает практика, использование специальных эвристических процедур, направленных на локальное улучшение решений по некоторому критерию, в составе генетического алгоритма позволяет повысить качество поиска в задачах с одним критерием [2]. Данное обстоятельство дает право сделать предположение о допустимости применения таких процедур и при решении задач многокритериальной оптимизации. С другой стороны, если такое улучшение проводится только по одному критерию, то, скорее всего, найдутся решения с низкими значениями других частных критериев, что не позволит адекватно определить область из оптимально-компромиссных решений.

Как показывает практика [2], в среднем более адекватное определение Парето-границы происходит в случае, когда частный критерий и глубина локальной адаптации выбираются случайным образом.

Операторы репродукции решений. В качестве репродукции новых решений можно воспользоваться схемами порядковых кроссоверов [3], [5]. Данные операторы манипулируют перестановками и из двух родительских перестановок получают две новые перестановки потомков. Данные операторы применяются только к первому уровню кодировок.

Многокритериальный отбор. Если в задачах с единственным критерием в результате отбора предпочтение отдавалось только лучшим особям, то в многокритериальном случае понятие «лучшие» следует заменить на «недоминируемые», т. е. следующее поколение составят, по возможности, только те особи, для которых не найдется решений, лучших по всем критериям.

Таким образом, оператор отбора должен формировать популяцию нового поколения, по возможности, только из недоминируемых особей с различными генотипами. Если таких недоминируемых решений недостаточно, тогда добирается популяция из особей, максимально близких к Парето-границе. Если же недоминируемых решений больше, чем нужно, среди этих кандидатов проводится дополнительный отбор, с тем чтобы новое поколение было максимально равномерно (в географическом смысле) распределено вдоль Парето-границы. Для того чтобы не потерять в процессе поиска недоминируемые решения, предлагается идея использовать хранилище для отбрасываемых решений, так как данные решения могут оказаться полезными.

Если при реализации процедуры отбора будет ощущаться нехватка недоминируемых решений, а в хранилище не будет ничего подходящего, то добирать популяцию придется из доминируемых решений, отдавая предпочтения тем, которые ближе к Парето-границе.

Для этого введем понятие *ранга* особи. Ранг особи $r(a_i)$ будем определять через количество особей a_i в текущей популяции, которые доминируют.

В приведенном примере (рис.2) “A”, “B”, “C” – недоминируемые решения бикритериальной задачи. Ранг “D” = 1, ранг “E” = 2, ранг “F” = 5. Таким образом, популяция нового поколения будет формироваться в первую очередь из особей “A”, “B” и “C”, затем из “D”, далее из “E”, но меньше всего шансов попасть в новую популяцию у особи “F”.

Итак, оператор отбора последовательно включает в популяцию сначала все недоминируемые особи, затем все особи первого ранга и т. д., до тех пор пока популяция не заполнится. Для предотвращения случая переполнения популяции предлагается алгоритм исключения “лишних” особей одного ранга, работа которого определяется следующими правилами: во-первых, необходимо удалять особи так, чтобы не нарушалось равномерное распределение особей вдоль Парето-границы; во-вторых, если особь удаляется из множества недоминируемых особей, то она заносится в хранилище. Рассмотрим данный алгоритм по шагам:

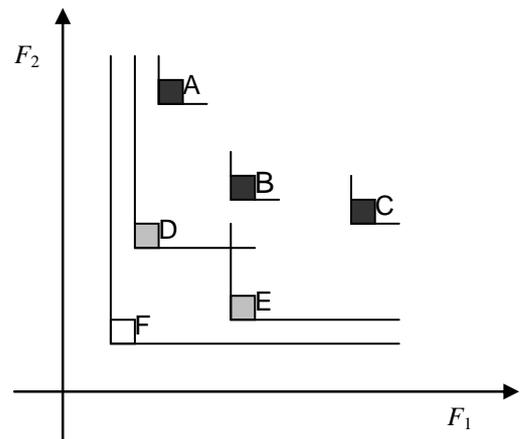


Рис. 2

1. $B(r)$ – множество решений ранга r .
2. N – число удаляемых решений ранга r .
3. $M = \|B(r)\| - N$ – число включаемых решений ранга r в популяцию.
4. Упорядочиваем в порядке возрастания критерия F_1 все решения из $B(r)$
5. Разбиваем $B(r)$ на локальные подмножества $B_i(r)$, удовлетворяющие условию $r(a_i, a_k) \leq \square_F$, где a_i – базовая точка из $B_i(r)$, $r(a_i, a_k) = \sqrt{\sum_j (F_j^i - F_j^k)^2}$ – евклидово расстояние между векторами приспособленностей особей a_i, a_k . Таким образом, разумно было бы подобрать \square_F с таким расчетом, чтобы число локальных групп $B_i(r)$ было равно M . Для этого предлагается процедура подбора радиуса \square_F и удаления из $B(r)$ лишних решений:

- а) \square_{\min} – минимальное расстояние между соседними точками из $B(r)$ в пространстве критериев;
- б) \square_{\max} – максимально возможный радиус – расстояние между крайними точками из $B(r)$;
- в) если $\square_{\min} = \square_{\max}$, то поиск не имеет смысла, случайным образом выбираем N особей и удаляем их из $B(r)$, выход;
- г) $\square_F = (\square_{\min} + \square_{\max}) : 2$;
- д) посчитать L , где L – число локальных групп с таким \square_F ;
- е) если $L > M$, то сделать $\square_{\max} = \square_F$. Переход к п. «г»;
- ж) если $L < M$, то сделать $\square_{\min} = \square_F$. Переход к п. «г».

6. Нужное \square_F найдено, из $B(r)$ удаляем все решения, которые не являются базовыми в подмножествах $B_i(r)$.

7. Выход.

Дополнительный механизм уточнения Парето-границы. В процессе многокритериального генетического поиска может быть построена область компромиссных решений, которая имеет разрывы и провалы (рис. 3). Это может быть следствием особенностей исследуемой задачи, а может быть обусловлено особенностями алгоритма и генетического поиска.

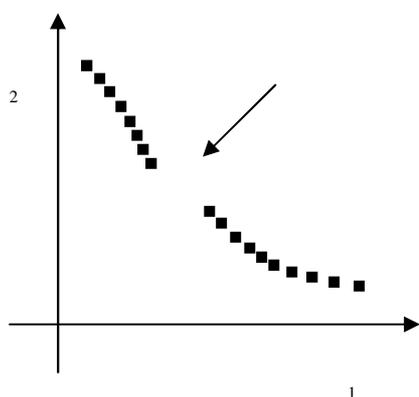


Рис. 3

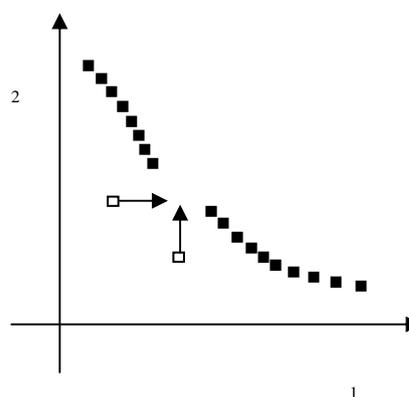


Рис. 4

Для «коррекции» области компромиссов предлагается использовать локальную адаптацию решений, находящихся вблизи (в географическом смысле в пространстве критериев) проблемных областей (рис. 4).

СПИСОК ЛИТЕРАТУРЫ

1. Бершадский А. М. Применение графов и гиперграфов для автоматизации конструкторского проектирования РЭА и ЭВА. Саратов: Из-во Саратов. ун-та, 1983.
2. Батищев Д. И., Старостин Н. В. Гибридный подход к решению экстремальных задач на графовых структурах // Изв. СПбГЭТУ "ЛЭТИ". Сер. "Информатика, управление и компьютерные технологии". 2002. Вып.3. С. 10-17.
3. Батищев Д. И., Старостин Н. В., Дроздова Е. П. Экстремальные задачи правильной раскраски графа// Сб. науч. тр. / ВПИ. Воронеж, 2000. Ч. 2. С. 49-60.
4. Батищев Д. И. Генетические алгоритмы решения экстремальных задач: Учеб. пособие под ред. Я. Е Львовича. Воронеж, 1995.
- 5 Популяционно-генетический подход к решению задач покрытия множества: Учеб. Пособие / Д. И. Батищев, В. Е. Костюков, Н. В. Старостин, А. И. Смирнов. Нижний Новгород: Изд-во ННГУ, 2004.

N. V. Starostin

GENETIC ALGORITHM (GA) WITH COMBINED SOLUTION REPRESENTATION FOR MULTICRITERIA PARTITIONING PROBLEM

Partitioning problems which should be formulated as multicriteria weighted hypergraph decomposition problems are under consideration. Hybrid, GA-based solving scheme is suggested. Particular qualities of direct and serial solution representations are discussed. It is suggested using of combined symbolic model. The schemes of appreciation, refining and reproduction of solutions are given. The multicriteria selection mechanism is described.

Genetic algorithm, multicriteria weighted hypergraph decomposition problem

РЕШЕНИЕ ЗАДАЧИ ОПТИМИЗАЦИИ НЕСТАЦИОНАРНОЙ ФУНКЦИИ ПРИ ПОМОЩИ ГЕНЕТИЧЕСКОГО АЛГОРИТМА С ИСПОЛЬЗОВАНИЕМ БАЗЫ ОПЫТА

Рассматриваются методы явной работы с памятью, когда полученные решения для определенных состояний внешней среды запоминаются с целью дальнейшего их использования при наступлении сходных состояний в будущем.

Генетический алгоритм, оптимизация, нестационарная функция

Рассматривая классификацию существующих подходов к решению задачи оптимизации нестационарной функции, можно выделить 2 основных класса:

- методы, поддерживающие или искусственно увеличивающие генетическое разнообразие популяции для повышения адапционных свойств генетического алгоритма;
- методы, использующие память с целью «восстановления» решений при наступлении сходных параметров внешней среды.

Конечно, это общая классификация и существуют методы, которые можно отнести как к первому, так и ко второму классу (например, использование диплоидного представления), но, в целом, она отражает существующую ситуацию в этой области.

Остановимся подробнее на одном из методов второго класса. Эти методы можно также разделить по признаку явного и неявного использования памяти. К методам неявного использования памяти относятся диплоидное и мультиплоидное представление генотипа [6], [7], [12], а также использование структурного (многослойного) представления генотипа [3], [9]. Кроме того, существуют методы, явно использующие память. Основное различие этих методов состоит в стратегиях вовлечения ранее полученных решений в текущее поколение и включения текущих решений в память [2], [4], [5], [10].

Постановка задачи оптимизации нестационарной функции. Обычно при решении задач оптимизации объективная функция не зависит от времени. Если снять это ограничение, то критерий оптимальности $\mathfrak{Z}(x)$ будет представлять собой нестационарную среду, в которой с течением времени возможны изменения как вида оптимизируемой функции $\mathfrak{Z}(x,t)$, так и значений оптимальных решений.

Функцию $\mathfrak{Z}(x,t)$ можно представить в виде набора независимых функций $F_i(x)$, периодически сменяющих друг друга с течением времени. При этом изменения происходят дискретно и функции $F_i(x)$ независимы друг от друга:

$$\mathfrak{Z} = \begin{cases} F_1, & t \in [0, t_1], \\ F_2, & t \in [t_1, t_2], \\ \dots & \\ F_N, & t \in [t_{N-1}, t_N]. \end{cases}$$

При оптимизации в условиях изменяющейся внешней среды (а именно так можно рассматривать изменения оптимизируемой функции) основной задачей генетического алгоритма является отслеживание динамики оптимума по пространству поиска.

Если принять, что критерий оптимальности меняется периодически, то наступит время, когда состояние внешней среды повторится, причем в прошлом генетический алгоритм уже находил для этого состояния приемлемые решения. Таким образом, идея запоминать успешные решения для некоторого состояния среды и использовать их при возвращении среды в это состояние кажется естественной.

Практическое использование памяти. Известный недостаток классического генетического алгоритма (ГА) – его быстрая сходимость в окрестности оптимума. При оптимизации нестационарной функции преждевременная сходимость является также нежелательной. Решая данную задачу ГА должен уметь достаточно быстро реагировать на изменение состояния внешней среды.

Процесс адаптации особей к нестационарной среде можно представить как некий циклический процесс: изменения условий внешней среды заставляют особи генетического алгоритма адаптироваться к новым условиям, что приводит к нахождению новых квазиоптимальных решений для текущего состояния (рис. 1). Однако при работе с нестационарной средой возникает одна и та же проблема: процесс адаптации происходит настолько успешно, что особи, соответствовавшие квазиоптимальным решениям при предыдущих состояниях среды, теряются. Таким образом, популяция ГА сама по себе не может достаточно хорошо сохранять нужные решения, поэтому необходимо использовать различные его модификации.

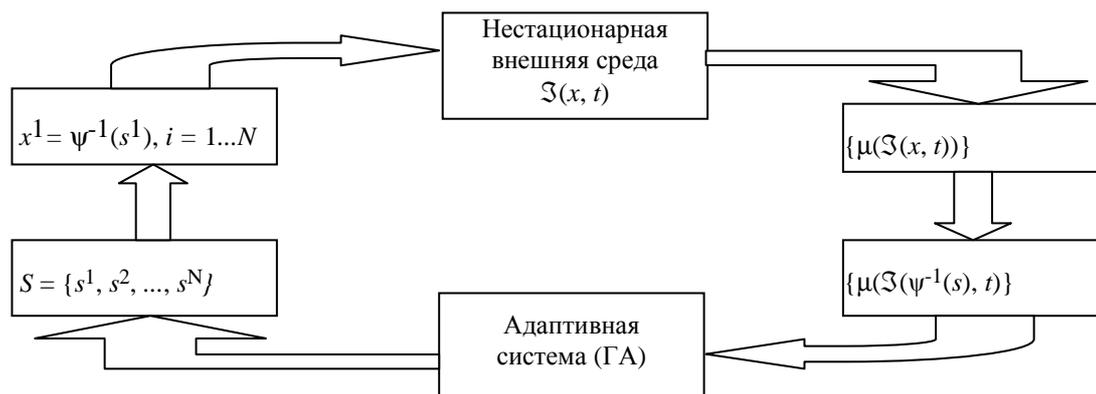


Рис. 1

Если известно, что внешняя среда меняется периодически, то полезно запоминать удачные решения в *базе опыта* при различных состояниях среды с целью их дальнейшего использования при повторении этого состояния в будущем.

Введение в ГА *базы опыта* является в некотором роде формой долговременного элитизма. Элитизм заставляет ГА сохранять лучший элемент каждого поколения, поэтому текущий оптимум никогда не будет потерян.

При использовании такого подхода возникает ряд вопросов:

- по какому принципу элементы будут добавляться в базу опыта?
- в каком случае они будут из нее исключаться?
- в каком виде элементы будут представлены в базе опыта?
- как в дальнейшем эти элементы будут использованы в решении?
- каков размер базы опыта?

Таким образом, ответы на эти вопросы определяют особенности конкретного метода. Рассмотрим некоторые из них.

Интересная стратегия рассматривается в работе [8], где лучшая особь в каждом поколении заносится в память, а другая особь удаляется из памяти в зависимости от ее возраста и вклада ее генотипа в разнообразие популяции в памяти. Особи из памяти затем используются в качестве потенциальных кандидатов в процессе естественного отбора (в дополнение к обычной популяции). Однако на практике этот метод не был опробован.

В своей работе Branke [2] сравнивает несколько стратегий замещения и внесения новых особей в память. Показано, что простое замещение наиболее близких особей дает ту же эффективность, что и замещение худшего из двух наиболее близких особей в памяти и обе эти стратегии работают лучше, чем схема замещения с целью увеличения разнообразия в памяти.

Использование базы опыта для прогноза вида лучших особей приведено в [4], [5]. Лучшая особь из текущего поколения добавляется в память, как только обнаружено изменение внешней среды, при условии, что ее генетический материал уже не содержится в памяти. Для использования особей из памяти в текущей популяции вся память оценивается при данных условиях внешней среды и из нее берется лучшая особь. Механизм предсказания основан на измерениях фенотипа особей в памяти при различных состояниях внешней среды.

Наиболее близким к описываемому в данной статье оказался метод явной работы с памятью у Ramsey и Grefenstette [10]. Они используют базу знаний для сохранения преуспевающих особей в долговременной памяти. При этом предполагается, что состояние среды может быть измерено. Через определенные промежутки времени лучшие особи сохраняются в базе знаний и индексируются вместе с данными, характеризующими состояние среды на данный момент. Как только состояние внешней среды меняется (меняются характеристики среды), алгоритм запускается заново. Для перезапуска половина начальной популяции берется из числа особей из базы знаний, которые были успешными при сходном состоянии внешней среды.

Узким местом данного алгоритма является возможность измерить и численно выразить состояние внешней среды (очевидно, что это возможно далеко не для всех задач).

Применение алгоритма с памятью для решения нестационарной задачи о 0-1 ранце. Задача о 0-1 ранце с периодическим изменением максимально допустимого суммарного веса стала классической тестовой задачей [4]–[6]. Стационарная задача о ранце формулируется следующим образом: *из заданного набора предметов, каждый из которых имеет вес и стоимость, нужно выбрать несколько и наполнить ими ранец таким образом, чтобы суммарная выгода по выбранным предметам была наибольшей, а их общий вес не превосходил вместимость ранца:*

$$\left\{ \begin{array}{l} \max \sum_{i=1}^N x_i v_i, \\ \sum_{i=1}^N x_i w_i \leq W_{\max}, \\ x_i \in \{0, 1\}, i = 1, 2, \dots, N, \end{array} \right.$$

где: N – количество предметов (размерность задачи); v_i – ценность предмета; w_i – вес предмета; W_{\max} – вместимость ранца; x_i принимает значение 1, если предмет положен в ранец, и 0 в обратном случае.

Очевидно, что значения $\{v_i, w_i, W_{\max}\}$ являются параметрами задачи. Если вместимость ранца W_{\max} периодически меняется, то задача становится нестационарной.

При исследовании алгоритма рассматривались 2 варианта постановки: максимальный вес периодически изменялся между двумя и между четырьмя значениями. В качестве тестовой рассматривалась задача из [6] при весовых ограничениях W_{\max} : $W_1 = 100, W_2 = 60, W_3 = 34, W_4 = 20$.

Так как известно, что количество целевых функций мало, то эффективным представляется сохранение текущей популяции при смене состояния внешней среды и возобновление поиска на данной популяции при повторении этого состояния.

В памяти сохраняется популяция и ее индикатор (в данном случае W_{\max} , при котором популяция работала). Индикатор несет исчерпывающую информацию о состоянии среды для успешного сохранения и извлечения популяции из памяти.

В качестве исходного использовался ГА, основанный на гаплоидном представлении особей. Для его улучшения были добавлены методы генерации начальной популяции с использованием эвристических подходов. Основная схема ГА оставалась без изменений, также были использованы классические виды кроссовера и мутации (одноточечные).

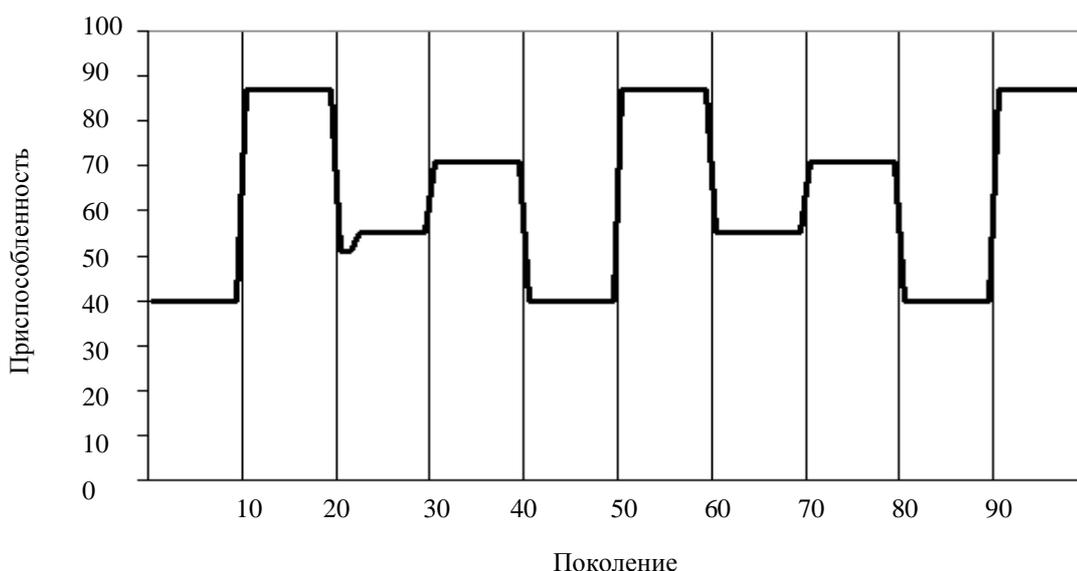


Рис. 2

Описанный метод показал отличную эффективность как в случае с двумя, так и с четырьмя периодически изменяющимися ограничениями. На рис. 2 приведен график отражающий динамику приспособленности при четырех весовых ограничениях. Хорошо виден ступенчатый характер функции приспособленности.

СПИСОК ЛИТЕРАТУРЫ

1. Branke J. Evolutionary approaches to dynamic optimization problems – Updated survey. GECCO Workshop on Evolutionary Algorithms for Dynamic Optimization Problems. 2001. P. 27-30.
2. Branke J. Memory enhanced evolutionary algorithms for changing optimization problems // Congress on Evolutionary Computation CEC99, 1999.
3. Dasgupta D. Optimization in Time-Varying environments using Structured Genetic Algorithms. Technical Report IKBS17-93 / University of Strathclyde, Dept. of Computer Science, 1993.
4. Raising the dead; extending evolutionary algorithms with case-based memory / J. Eggermont, T. Laerts, S. Poyhonen, A. Termier // 2nd European Conf. Of Genetic Programming (EuroGP'01), 1997. Vol. 2038. P. 350–361

5. Eggermont J., Lenaerts T. Non-stationary function optimization using evolutionary algorithms with a case-based memory. Technical Report TR2001 / University of Strathclyde, Dept. of Computer Science, 2001.
6. Goldberg D. E., Smith R. E. Nonstationary function optimization using genetic algorithm with dominance and diploidy // 2nd Intl. Conf. of Genetic Algorithms, Lawrence Erlbaum Associates. 1987. P. 59-68
7. Hadad B. S., Eick C. F. Supporting polyploidy in genetic algorithms using dominance vectors // 6th Intl. Conf. of Evolutionary Programming. 1997. Vol. 1213. P. 223-234.
8. Adaptation to changing environments by means of the memory based thermodynamical genetic algorithm / N. Mori, S. Imanishi, H. Kita, and Y. Nishikawa // Intern. Conf. on Genetic Algorithms, Morgan Kaufmann Publishers. 1997.
9. Ohkura K., Ueda K. Structured String Representation and Adaptive Genetic Search // Japan/USA Symposium on Flexible Automation. 1996. Vol. 2. P. 1477-1480.
10. Ramsey C. L., Grefenstette J. J. Case-based initialization of genetic algorithms // 5th 2nd Intl. Conf. of Genetic Algorithms, 1993. P. 84-91.
11. Ryan C. Diploid without dominance // Third Nordic Workshop on Genetic Algorithms, 1997. P. 63-70.
12. Smith R.E. Diploid genetic algorithms for search in time varying environments // In Annual Southeast Regional Conference of the ACM, New York, 1987. P. 175-179.

D. I. Batishev, E. A. Neumark

NONSTATIONARY FUNCTION OPTIMIZATION USING GENETIC ALGORITHM WITH CASE-BASED MEMORY

In this paper the methods which use memory explicitly are observed. Obtained solutions are stored in case-based memory in order to use them when the same environment occurs.

Genetic algorithm, optimization, nonstationary function

УДК 621.372.001.24:681.3

В. А. Павлушин, Ф. А. Михеев, М. В. Марков

**ГЕНЕТИЧЕСКИЕ АЛГОРИТМЫ РЕШЕНИЯ ЗАДАЧ
ОПТИМИЗАЦИИ ЧАСТОТНЫХ ХАРАКТЕРИСТИК
РАДИОЭЛЕКТРОННЫХ СХЕМ**

Рассматриваются популяционно-генетические алгоритмы применительно к решению задач оптимизации частотных характеристик.

Оптимизация частотных характеристик, генетический алгоритм, САПР оптимального проектирования РЭС

Задача оптимизации частотных характеристик радиоэлектронных схем хорошо формализована и достаточно успешно решается [1]. Однако есть ряд проблем, связанных с выбором алгоритма оптимизации. Этот алгоритм должен учитывать прямые ограничения на вектор варьируемых параметров, быть устойчивым к сложному характеру целевой функции и иметь глобальные свойства. В данной статье рассматривается такой алгоритм, основанный на эволюционно-генетическом подходе.

В основе типового генетического алгоритма лежит идея эволюционного развития популяции особей [2]. Предполагается, что существует группа особей – популяция. Каждая особь содержит набор генов. При размножении потомки наследуют гены от родителей. Возможны мутации, при которых гены особи изменяются случайным образом. В процессе естественного отбора в популяции остаются лишь особи с наибольшей приспособленностью, остальные исключаются из эволюционного развития. Схема типового генетического алгоритма представлена на рис. 1.

Для оптимизации частотных характеристик эта схема применяется следующим образом. Каждая особь соответствует определенной точке в области допустимых решений. Ка-

ждая координата вектора варьируемых параметров соответствует одному гену. Цель эволюции – найти наиболее приспособленную особь, которой соответствует минимальное значение целевой функции. При оптимизации частотных характеристик в качестве целевой функции принимают среднеквадратическое отклонение от желаемой характеристики:

$$Q(x) \rightarrow \min,$$

$$Q(x) = (Y_{\text{жел}}(x) - Y(x))^2,$$

$$a \leq x \leq b,$$

$$x, a, b \in R^n,$$

$$F(x) = \frac{1}{Q(x)^2 + 1},$$

где $Q(x)$ – целевая функция, принимающая только неотрицательные значения; x – вектор варьируемых параметров схем; $Y_{\text{жел}}(x)$ – желаемая характеристика; $Y(x)$ – частотная характеристика схемы; a, b – верхняя и нижняя границы изменения вектора x ; $F(x)$ – функция приспособленности.

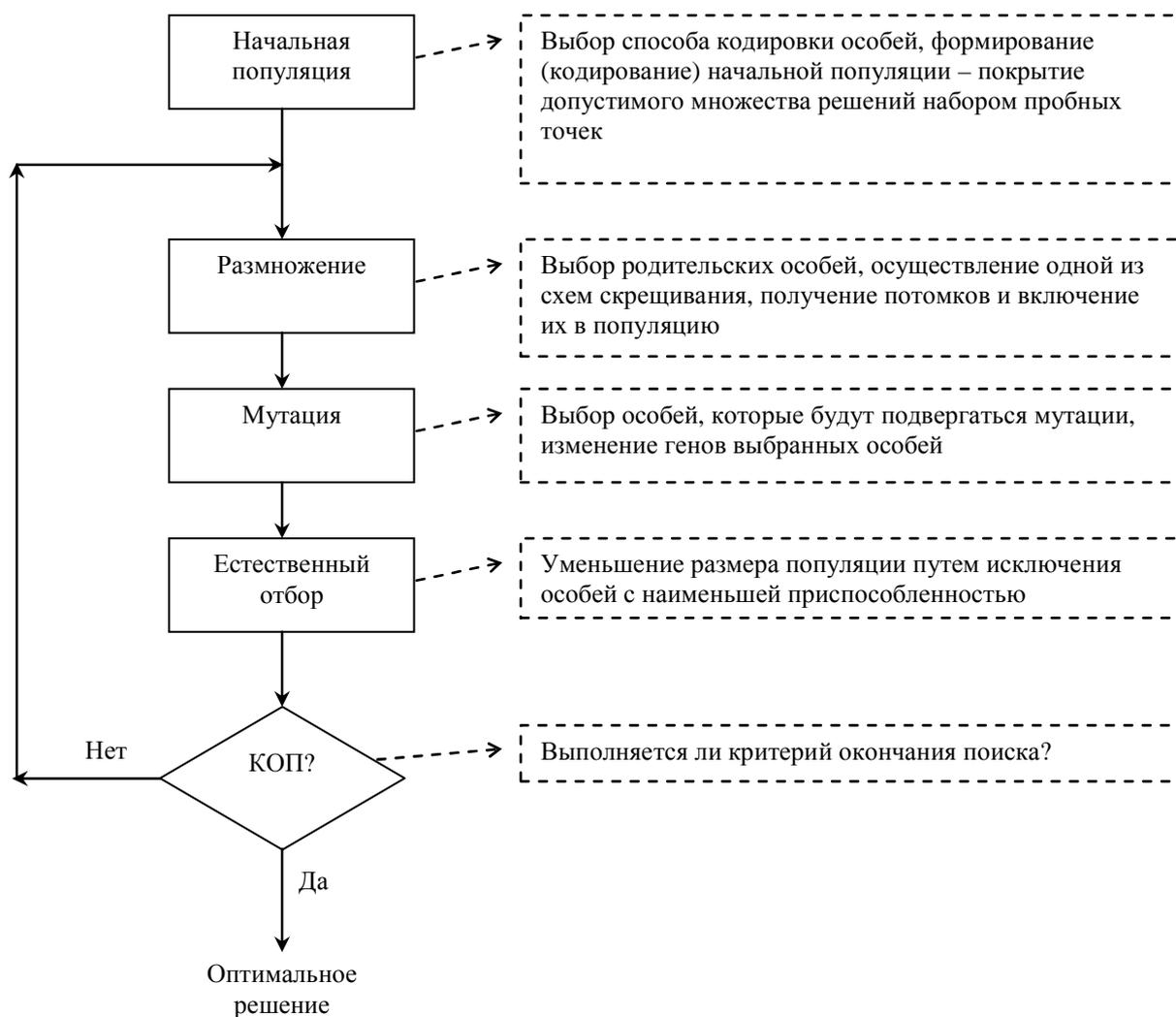


Рис. 1

Для установления соответствия особей популяции и точек допустимой области используют один из двух способов кодирования: вещественный или целочисленный (двоичный).

При вещественном способе кодирования ген особи численно равен соответствующей координате точки допустимой области.

При целочисленном, или двоичном кодировании i -й ген представлен как набор бит, которые задают целое число k_i . По нему определяют вещественную координату точки следующим образом:

$$x_i = a_i + k_i \frac{(b_i - a_i)}{m_i}, \quad (1)$$

где a_i, b_i – нижняя и верхняя границы переменной x_i ; k_i – код i -го гена; m_i – число точек разбиения по i -й координате. Код гена k_i лежит в пределах от 0 до m_i . Число m_i выбирается так, чтобы выполнялось условие

$$\frac{(b_i - a_i)}{m_i} \geq \varepsilon, \quad (2)$$

где ε – точность поиска оптимума. Весь набор генов $i = 1..n$ объединяют в одну битовую цепочку.

| | | | | | |
|-------|-------|---------|-------|---------|-------|
| k_1 | k_2 | \dots | k_i | \dots | k_n |
|-------|-------|---------|-------|---------|-------|

Способ кодирования особей выбирают на начальном этапе, и он влияет на все последующие шаги генетического алгоритма. Для формирования начальной популяции допустимую область решений покрывают набором точек, число которых выбирается заранее и называется численностью популяции M . Как правило, точки начальной популяции получают с помощью генератора случайных чисел. Далее генетический алгоритм входит в цикл эволюционного развития.

Операция размножения направлена на увеличение численности популяции. Для размножения из популяции случайным образом выбирают две родительские особи. Затем производят обмен генов между родителями и получают наборы генов для потомков, которые включаются в популяцию. Обмен генов чаще всего проводят путем одноточечного кроссинговера. Операцию размножения повторяют, пока численность популяции M не удвоится.

Операцию мутации производят для внесения в популяцию новой генетической информации, которой нет ни в родителях, ни в потомках. Для каждой особи с вероятностью мутации P_m изменяют значение кода случайно выбранного гена.

Операция естественного отбора направлена на уменьшение численности популяции до исходных размеров M . Из популяции исключаются особи с наименьшей приспособленностью. Для задачи на поиск минимума это соответствует исключению точек, значение целевой функции в которых больше, чем в точках, остающихся в популяции.

В качестве критерия окончания поиска обычно используют достижение популяцией определенного поколения, т. е. выполнение заданного числа шагов. Развитие популяции направлено на увеличение средней приспособленности по всей популяции:

$$F_{\text{ср}} = \sum_{i=1}^M F(x^i),$$

поэтому можно утверждать, что после работы генетического алгоритма будут найдены точки значение целевой функции в которых лучше, чем в точках исходной популяции. Поиск можно прекратить и раньше в случае вырождения популяции.

Популяция считается вырожденной, если приспособленность по всей популяции примерно одинакова: $\sqrt{\sum_{i=1}^M (F(x^i) - F_{cp})^2} \leq \varepsilon$.

Генетический алгоритм в чистом виде по своим вычислительным свойствам похож на алгоритм случайного поиска. Для улучшения алгоритма существуют различные модификации на каждом его этапе.

Как показывает практика кодирование генов вещественными числами проще в реализации, однако дает несколько худшие результаты, чем целочисленное кодирование. Связано это с тем, что при целочисленном кодировании обычно используют код Грея. При вещественном кодировании на этапе мутации весь ген изменяют на случайное число, а при двоичном кодировании изменяют один бит гена, что для кода Грея соответствует изменению на 1 целочисленного представления гена. Таким образом, соответствующая вещественная координата изменяется не на случайное число, а на другое близкое к начальному число. В операции размножения при двоичном кодировании возможно разбиение набора генов для кроссинговера не по границе генов, а по границе бит в генах.

При двоичном кодировании для поиска оптимального решения с заданной точностью число точек разбиения допустимого интервала по каждой координате m_i определяется из условия (2). При точном поиске на большом интервале можно получить достаточно большое число m_i – для реальных задач более 10^7 . С ростом числа точек разбиения увеличивается число бит, кодирующих ген, и возрастает число шагов, необходимых для того, чтобы популяция сошлась. В связи с этим целесообразно использовать лишь глобальные свойства генетического алгоритма и применять крупную сетку разбиения, а на этапе уточнения решения применять методы на основе производных. Для задачи оптимизации частотных характеристик используют логарифмическую шкалу изменения варьируемых параметров. В этом случае выражение (1) принимает вид

$$x_i = 10^{\log a_i + \frac{k_i}{m_i} (\log b_i - \log a_i)}.$$

Для улучшения равномерности изначального покрытия допустимой области, целесообразно использовать точки LPτ -последовательности [3].

В целях повышения сходимости популяции в качестве родительских особей для размножения целесообразно выбирать наиболее приспособленные. Для этого используют рулеточный отбор, при котором вероятность особи стать родителем определяется колесом рулетки, а размер сектора соответствующий особи пропорционален ее приспособленности.

Для предотвращения преждевременного вырождения популяции и остановки в неоптимальной точке вероятность мутации должна быть не меньше 3...5%. При этом совместно с точечной мутацией используют макромутацию, при которой изменению подвергается не только один бит в гене, но и целая последовательность бит в битовом наборе, кодирующем цепочку генов.

Наиболее удачной модификацией является генетический алгоритм с градиентным обучением лидера. Суть его в следующем. В популяции выделяется лидер – особь с наибольшим значением приспособленности. Лидер обязательно проходит в следующую популяцию. Для лидера запускается алгоритм на основе производной с проекцией градиента на

допустимую область. Затем получают двоичную кодировку «улучшенного» лидера и возвращают его в популяцию. В таком варианте алгоритм приобретает хорошую сходимость в области притяжения минимума, свойственную методам на основе производных, и в то же время обладает глобальной сходимостью и робастностью по отношению к виду целевой функции, что характерно для методов нулевого порядка. Метод оперирует непосредственно множеством допустимых решений, учитывая, таким образом, прямые ограничения.

Исходя из всего изложенного, представляется целесообразным включение описанного метода в подсистему оптимизации диалоговой САПР оптимального проектирования радиоэлектронных схем [4].

Рассмотрим пример практического применения предложенного алгоритма. Задана принципиальная схема фильтра нижних частот (рис. 2).

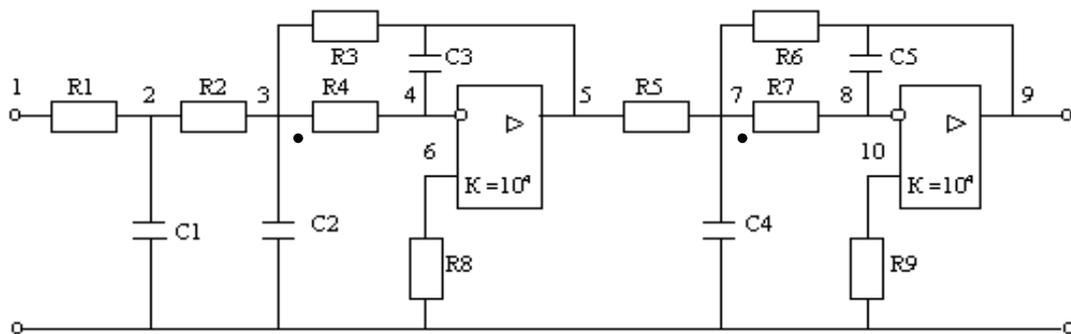


Рис. 2

Необходимо обеспечить заданную частотную характеристику путем варьирования параметров компонентов $C1$, $C2$, $C4$, $R1$, $R6$. Желаемая характеристика: -6 дБ до частоты 1 кГц, менее -50 дБ после частоты 2 кГц.

Векторы ограничений на варьируемые параметры:

$$\mathbf{a} = (0,01; 0,01; 0,01; 1; 1)^t, \quad \mathbf{b} = (1; 1; 1; 100; 100)^t.$$

Полученный оптимальный вектор $\mathbf{x}^* = (0,111; 0,264; 0,102; 10; 10)^t$, частотная характеристика до и после оптимизации, а также желаемая частотная характеристика представлены на рис. 3.

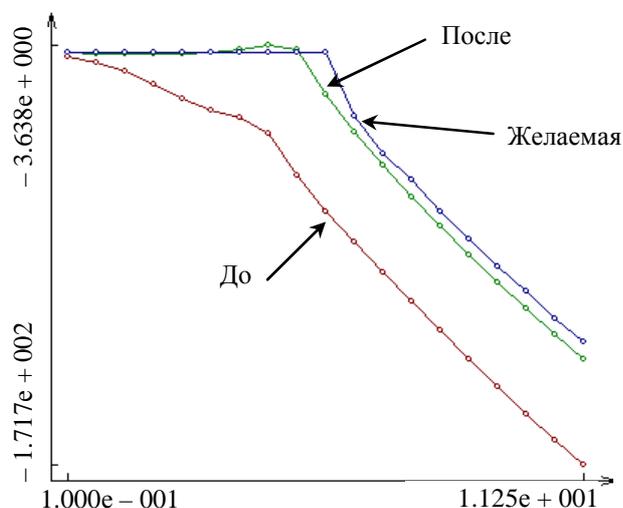


Рис. 3

В заключение отметим, что по сравнению с традиционным подходом предложенный алгоритм сокращает количество вычислений целевой функции. По результатам проведенных экспериментов число вычислений уменьшилось в 1,5 раза.

СПИСОК ЛИТЕРАТУРЫ

1. Аль-Русан М. Б. Х., Павлушин В. А. Программа параметрической оптимизации радиоэлектронных схем // Изв. СПб ГЭТУ «ЛЭТИ» серия «Информатика, управление и компьютерные технологии». 2002. №1. С. 7–8.
2. Батищев Д. И. Генетические алгоритмы решения экстремальных задач: Учеб. пособие / Под ред. Львовича Я.Е., Нижний Новгород, Изд-во Нижегородского государственного университета имени Н. И. Лобачевского 1995.
3. Соболев И. М., Статников Р. Б. Выбор оптимальных параметров в задачах с многими критериями. М.: Наука, 1981.
4. Дмитриевич Г. Д., Павлушин В. А. Учебная подсистема оптимального проектирования для САПР РЭС // Междунар. конф. современных технологий обучения, СПб. 20 апреля 2005: том 2. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2005, Т.2. С.108–109.

V. A. Pavlushin, P. A. Miheev, M. V. Markov

GENETIC ALGORITHM IN FREQUENCY CHARACTERISTICS OPTIMIZATION

The approach to optimization of frequency characteristics based on genetic algorithm has been presented.

Genetic algorithm, electronic circuits design, optimization

УДК 514.7.1

С. Г. Иванов, С. Н. Поздняков

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ФОРМИРОВАНИЯ МАТЕМАТИЧЕСКИХ ПОНЯТИЙ

Обсуждаются базовые структуры, необходимые для компьютерного моделирования обучения человека простейшим математическим понятиям.

Информационная среда, машинное обучение, диалоговые системы, компьютерная поддержка обучения

В последнее десятилетие окончательно сформировалось основное направление в развитии средств компьютерной поддержки обучения фундаментальным наукам в школе. Оно связано с развитием инструментальных средств предметного моделирования. В физике это проявилось в создании различных физических лабораторий, в математике – в создании группы программных продуктов для геометрического моделирования (The Geometer's Sketchpad, Cabri, Cinderella и пр.) [1].

В предыдущих работах [2–5] авторами была показана возможность создания в рамках подобных инструментальных сред специальных манипуляторов, поддерживающих самостоятельность обучаемого и направляющих ее на овладение определенными математическими понятиями. Была показана возможность создания на базе манипуляторов специальных задачников, в которых осуществляется верификация частичных решений и поддерживается поисковая деятельность обучаемого в процессе решения задачи [6].

Следующим шагом в решении проблем, связанных с построением компьютерной поддержки обучения фундаментальным наукам, является, по мнению авторов, соединение инструментальных возможностей компьютера и средств самоконтроля деятельности обу-

чаемого с механизмами машинного обучения [7–9]. Введение элементов машинного обучения в учебно-методические программные комплексы позволит изменить психологическую структуру процесса обучения, заменив пассивную функцию обучаемого активной функцией обучающего (на практике будут сочетаться обе функции).

Для создания таких систем необходимы более тонкие механизмы внутримашинного представления действий, которые совершает человек с объектами информационной среды, определяемой компьютерными моделями предметной области.

В данной статье выделяются и анализируются структурные компоненты среды, в которой моделируется формирование математических понятий. Предложенные структуры иллюстрируются на примере начального этапа формирования понятия натурального числа.

Мир шаров на плоскости как модель множества однородных объектов. Построим модель формирования простейших математических понятий, основываясь на множестве плоских однородных геометрических объектов – шаров на плоскости.

Для того чтобы система была активной, она должна иметь возможность воздействия на шары, в данном случае – сдвигать шары на небольшие расстояния в различных направлениях.

Далее, система должна обладать возможностью выделять характеристики расположения шаров (элементы модели зрения), выделять группы шаров (элементы модели внимания), давать конфигурациям внутренние имена и обращаться к ним по имени (элементы модели памяти), сравнивать конфигурации и образовывать понятия (элементы модели интеллекта).

В целом все компоненты модели можно подразделить на две группы:

- компоненты, моделирующие внешнюю среду (в данном случае это «мир шаров»);
- компоненты, моделирующие внутренние характеристики системы, включающие специальные механизмы, необходимые для ее обучения.

Первая группа, связанная с моделированием предметной среды, исследована авторами в предыдущих работах.

Предметом настоящей статьи является моделирование внутренних характеристик системы. Выбрав для иллюстрации простой пример, авторы не стремились показать все многообразие проявлений этих компонентов в различных и более содержательных предметных моделях, а лишь выделить базовые типы этих компонентов.

«Генетические компоненты». Рассматривая процесс моделирования обучения как совокупность большого числа различных предметных моделей, отвлечемся пока от изучения вопроса о преемственности этих моделей, от того, как последовательно формируются *все* используемые человеком понятия. Для этого мы введем термин «генетические компоненты», которые считаются для данной модели уже существующими и работающими по описанным правилам.

Пример моделирования «зрения».

Будем основываться на некотором множестве операций обработки информации о взаимном расположении шаров, которые существуют у исходной обучаемой системы «генетически», т. е. которым обучить невозможно.

Положим, что «глаз» создаваемой системы характеризуется следующими операциями:

- 1) он может определять, образует ли данное подмножество шаров группу; например, соответствующий булевский признак S принимает значение «истина», если наибольшее расстояние между центрами выделенного подмножества шаров меньше наименьшего расстояния между элементами множества и его дополнения;

2) он может строить минимальное остовное дерево полного графа, построенного соединением центров шаров, если весом ребра считать расстояние между центрами;

3) он может определять признаки этого остовного дерева, например, признак W – является ли дерево путем, т. е. верно ли, что степени всех вершин не превышают 2.

Моделирование интеллектуальных функций.

1. Модель внимания

Отдельно рассмотрим произвольное (спонтанное) внимание и произвольное, вызванное извне или самой системой. Внимание будет определяться выделением объектов (как элементов внешней среды, так и элементов внутренней среды понятий – имен, формируемых в процессе обучения). Произвольное внимание связано со случайным выбором элементов, «привлечение внимания» действием обучающего, – их явным указанием, самопроизвольное внимание – выбором тех или иных элементов на основе сформированных понятий.

Пример моделирования внимания.

В «мире шаров» произвольное внимание будет случайно выделять группу шаров, попадающих в сферу случайного радиуса R , центр которой спонтанно перемещается в некоторой области плоскости. Привлечение внимания определяется действием обучающего, перехватывающего управление «сферой внимания». Самопроизвольное внимание – выбор центра и радиуса сферы самой системой на основе внутренней информации.

2. Модель времени и временной памяти.

Для обучения необходим анализ временной динамики состояний. Для этого необходима модель времени и памяти. Будем считать время дискретным. В очередной момент времени во временной памяти фиксируется состояние системы, т. е. те значения параметров, которые соответствуют частям системы «под вниманием». Ограниченность памяти можно моделировать «затуханием» их со временем, т. е. исчезновением через некоторое количество тактов времени. Для начала можно взять модель, когда рассматриваются только 2 последних момента времени.

3. Модель активности.

Для того чтобы обучаться, система должна быть активной, ибо проверить обученность системы можно только на практике, т. е. ее действиями. Активность системы будем моделировать:

1) наличием движителя; движение объектов связано с перемещением зоны внимания: двигаться могут только объекты в зоне внимания, а сама зона внимания перемещается за передвигаемым объектом;

2) спонтанным переключением внимания с объекта на объект (это могут быть объекты внешней среды либо понятия – имена, данные учителем или системой);

3) реакцией на возникающие комбинации на «внешней» и «внутренней» сценах;

4) воспроизведением выученных действий, связанных с содержанием понятий;

5) формулировкой вопросов «учителю».

Пример.

В «мире шаров» система может передвигать шары по плоскости в восьми различных направлениях (север, северо-восток и т. д.) на небольшое фиксированное расстояние; естественно направление считать не абсолютным, а относительным – относительно текущего центра внимания; для обратной связи нужен параметр, характеризующий движение – до-

бавим булевский параметр C , соответствующий динамике процесса (приближаемся ли мы к зоне внимания предыдущего шага или нет). Система стремится делать так, чтобы C было истинным, т. е. в ней заложено стремление «собирать шары в кучу».

4. Модель формирования внутренних образов внешних объектов – именованная объектов

Это наиболее сложная и важная часть комплексной модели. Можно сказать, что работа ведется и на другом множестве объектов – множестве имен, которое устроено аналогично множеству объектов внешней среды и элементами которого являются символы, связанные с новыми понятиями. На это множество будут переноситься генетические механизмы распознавания и внимания. Именоваться будут также комбинации внешних объектов и символов, находящихся в зоне внимания.

Пример.

В мире шаров можно указать подмножество шаров, образующее группу, и назвать это понятие «вместе», или «группа», или «рядом».

Другой пример: 2 шара образовали «группу», тогда обязательно будет истинным и значение признака W – образуют путь. Можно назвать эту ситуацию «линия».

5. Модель озарения

Этот компонент является главной частью, обеспечивающей разумную самостоятельность системы. Если в процессе спонтанной или самоинициированной деятельности обнаруживается «узнаваемая» системой комбинация (т. е. именованная ранее), то происходит «озарение» и переключение внимания на возникшее понятие.

Пример.

В мире шаров спонтанное движение шаров может привести к тому, что попавшие в зону внимания окажутся «рядом», что система «поймет» за счет механизма озарения.

6. Модель формального обобщения и вопросов системы к учителю

Для описания внутренней активности системы нужен еще один механизм – механизм отбрасывания «ненужных» признаков и попытка идентифицировать ситуацию по урезанному множеству признаков. Поскольку эта операция может дать неверный результат, правильно ее трактовать как вопрос системы вида «Верно ли, что...?».

Пример.

В мире шаров 3 шара образовали «группу» и «путь». В мире имен они совпали, но остовные деревья у них разные. Система должна спросить «Это линия?». Если система получит положительный ответ (подкрепление), то понятие «линия» уже не будет связано с остовным деревом, а будет определяться признаком SW .

Перейдем к описанию работы системы и ее обучения.

Пример обучения системы порядковым числительным. Поместим в зону внимания 2 шара. На озарение по признаку SW и конкретному остовному дереву (состоящему из двух вершин и одного ребра) дадим название «два». Затем либо специальным «обучающим» воздействием, либо ожиданием спонтанного события придем к ситуации, когда в зоне внимания окажутся 3 шара на линии (признак SW и остовное дерево – путь на трех вершинах). На «озарение» системы дадим имя «три» (если этого не сделать, система будет пользоваться внутренним именем). Так можно продолжать именовать конфигурации и дальше, но утверждать, что понятие порядкового числительного сформировано, можно, если система освоит алгоритм построения следующего по предыдущему.

В формировании этого понятия будут играть роль более сложные механизмы, связанные с сохранением комбинаций последовательных конфигураций во времени (напомним, что в данной модели время дискретное и хранятся две последние конфигурации). Когда создавалось понятие «три», предыдущей временной конфигурацией было «два», следующей конфигурацией было *CSW* (движение шара в сторону двойки с последующим озарением на появление новой конфигурации *SW*). Таким образом, вместе с формированием понятия «три», появилось внутреннее понятие, состоящее в последовательном осуществлении *SW*, *CSW*. Вот это понятие и следует сопоставить понятию порядкового числительного.

Из приведенного анализа и модельного примера видно, что в структуру модели формирования математических понятий должны входить следующие компоненты:

- компоненты, моделирующие внешнюю среду;
- компоненты, моделирующие внутренние образы объектов внешней среды (имена);
- генетические механизмы анализа конфигураций объектов внешнего мира;
- компоненты, определяющие познавательную активность системы;
- компоненты, моделирующие выделение внешних объектов среды и внутренних образов (модель внимания);
- компоненты сопоставления объектов, находящихся в зоне внимания, внутренним объектам (модель озарения);
- компоненты, связанные с формированием абстрактных понятий и обобщений (модель формальных обобщений и вопросов системы к «учителю»).

СПИСОК ЛИТЕРАТУРЫ

1. Kortenkamp U. Foundations of Dynamic Geometry, Dissertation, ETH Zürich, Institut für Theoretische Informatik, November 1999.
2. Иванов С.Г., Поздняков С.Н. Компьютер в продуктивном обучении математике или как информационные технологии могут поддержать интеллектуальную свободу обучаемого. Компьютерные инструменты в образовании, 2003 г., № 5, с. 5–15.
3. Pozdnyakov S., Ivanov S. Computers in productive teaching of mathematics or how information technologies can support intellectual freedom of the learner. The 10-th Intern. Congress on Mathematical Education, National presentation: Russia, Selected materials, Copenhagen, Denmark, July 4–11, 2004. p. 115–124.
4. Pozdnyakov S. Mathematical models of intellectual dialogue in Computer Aided Learning. The First World Conference on Twenty-First Century Mathematics. School of Mathematical Sciences, GC University, Lahore, Punjab (Pakistan), 2004.
5. Башмаков М.И., Поздняков С.Н., Резник Н. А. Информационная среда обучения. СПб: СВЕТ, 1997.
6. Иванов С. Г. Компьютерная поддержка решения математических задач как средство организации продуктивной деятельности учащихся. Автореферат диссертации на соискание ученой степени кандидата педагогических наук. М.: 2004.
7. Психология машинного зрения: Сб. статей / Под ред. П. Г. Уинстона. М.: Мир, 1978.
8. Уинстон П. Г. Искусственный интеллект. – М.: Мир, 1980.
9. Minsky M. The Society of Mind. Simon and Schuster. 1987.

S. G. Ivanov, S. N. Pozdnyakov

COMPUTER MODELING OF HUMAN UNDERSTANDING OF SIMPLE MATHEMATICAL CONCEPTS

Basic structures for computer modeling of human understanding of simple mathematical concepts are discussed.

Informational environment, machine learning, dialog systems, computer support of learning

ИСПОЛЬЗОВАНИЕ АЛГЕБРАИЧЕСКОГО ПОДХОДА ДЛЯ ПОСТРОЕНИЯ АЛГОРИТМОВ НА ГРАФАХ

Рассматриваются возможности использования алгебраического подхода к исследованию и построению алгоритмов на графах.

Абстрагирование, алгоритм, граф, гомоморфизм, замкнутое полукольцо, матричный алгоритм, моноид, путь графа, эпиморфизм

Алгебраический подход к исследованию алгоритмов на графах был впервые введен в [1] при рассмотрении задачи о нахождении путей и, в частности, транзитивного замыкания графа. Затем в [2] данная идея была использована не только для анализа временных характеристик, а как подход, позволяющий обобщить целый класс матричных алгоритмов, применяемых, как известно, для разреженных графов.

Очевидно, что всякая классификация (в частности, алгоритмов) предполагает абстрагирование, т. е. выделение существенных свойств исследуемых систем, а значит, видение их в упрощенном виде. Зачастую полученный таким образом класс является в некотором смысле эквивалентным классу систем иной, даже весьма отдаленной предметной области. Данный факт позволил А. Ахо рассматривать множество меток (весов) путей графа вместе с введенными над ними операциями как замкнутое полукольцо. Это позволяет распространить на рассматриваемую предметную область гомоморфизм и изоморфизм, фундаментальные понятия современной алгебры. В статье описана возможность использования свойств алгебраических систем для построения алгоритмов на графах.

Определение замкнутого полукольца. Рассмотрим аксиоматику замкнутого полукольца (см. [1] или с незначительными изменениями [2]).

Определение: *Замкнутым полукольцом* называется система $(S, \oplus, \otimes, \bar{0}, \bar{1})$, где S – множество элементов, а \oplus и \otimes – бинарные операции на S , обладающие следующими свойствами:

1) $(S, \oplus, \bar{0})$ – моноид, т. е. множество S замкнуто относительно \oplus ($a \oplus b \in S$) для всех a и b из S . Операция \oplus ассоциативна $(a \oplus b) \oplus c = a \oplus (b \oplus c)$, для всех a, b, c из S ; $\bar{0}$ – единичный элемент моноида, т. е. $a \oplus \bar{0} = \bar{0} \oplus a$ для всех a из S . Тройка $(S, \otimes, \bar{1})$ также является моноидом. Кроме того, полагаем, что $\bar{0}$ служит аннулятором, т. е. $a \otimes \bar{0} = \bar{0} \otimes a = \bar{0}$.²

2) Операция \oplus коммутативна ($a \oplus b = b \oplus a$) и идемпотентна ($a \oplus a = a$).

3) Операция \otimes дистрибутивна относительно \oplus : $(a \oplus b) \otimes c = a \otimes c \oplus b \otimes c$ и $a \otimes (b \oplus c) = a \otimes b \oplus a \otimes c$.

² Хотя соотношение и вводится как аксиома, оно может быть доказано от противного, отталкиваясь от дистрибутивности.

4) Если $a_1, a_2, \dots, a_i, \dots$ – счетная последовательность элементов из S , то сумма $a_1 \oplus a_2 \dots \oplus a_i \dots$ существует и единственна. Более того, ассоциативность, коммутативность и идемпотентность выполняются не только для конечных, но и для бесконечных сумм.

5) Операция \otimes дистрибутивна не только относительно конечных сумм, но и относительно счетных (это не следует из свойства 3). Таким образом, из свойств 4 и 5 вытекает, что

$$\bigoplus_i a_i \otimes \bigoplus_j b_j = \bigoplus_{i,j} (a_i \otimes b_j)$$

Следуя далее [1], определим семантику полукольца применительно к графам. Допустим, (V, E) , $E \subseteq V \times V$ – некоторый орграф, в котором каждое ребро помечено элементом некоторого замкнутого полукольца $(S, \oplus, \otimes, \bar{0}, \bar{1})$. Определим метку пути как произведение меток ребер. Со- поставим *нуль-маршруту* (т. е. маршруту, не содержащему ребер) элемент $\bar{1}$. Для каждой пары вершин (i, j) определим $c(i, j)$ как сумму меток всех путей из i в j . Условимся считать, что сумма по пустому множеству путей равна $\bar{0}$.

Более формально, если λ – *функция разметки*

$$\lambda : V \times V \rightarrow S, \quad (1)$$

то для любых двух путей p_1 и p_2 верны следующие соотношения:

$$\begin{aligned} \lambda(p_1 * p_2) &= \lambda(p_1) \otimes \lambda(p_2), \\ \lambda(p_1 + p_2) &= \lambda(p_1) \oplus \lambda(p_2), \end{aligned} \quad (2)$$

где $*$ – операция *соединения* (конкатенации) путей $p_1 = \langle i, i+1, \dots, k \rangle$ и $p_2 = \langle k, k+1, \dots, j \rangle$:

$$p_1 * p_2 = \langle i, i+1, \dots, k, k+1, \dots, j \rangle, \quad (3)$$

дающая результат, отличный от *нуль-маршрута* (обозначим его как Θ) всякий раз, когда конец первого пути совпадает с началом второго; операция $+$ – операция *объединения* двух путей $p_1 = \langle i, i+1, \dots, k \rangle$ и $p_2 = \langle i', i'+1, \dots, j \rangle$:

$$p_1 + p_2 = \{ p_1, p_2 \}, \quad (4)$$

дающая в результате *семейство путей* между двумя конкретными вершинами i и j . В случае не- совпадения хотя бы одного из концов p_1 и p_2 результатом будет $\{ \}$ – пустое семейство путей.

Можно заметить, что операции $*$ и $+$ дистрибутивны, а именно

$$p_1 * (p_2 + p_3) = p_1 * p_2 + p_1 * p_3 = \{ p_1 * p_2, p_1 * p_3 \}, \quad (5)$$

причем операцию $*$ можно распространить и на семейства путей между вершинами. На- пример, для двухэлементных семейств можно написать следующее соотношение:

$$\{ p_1, p_2 \} * \{ p_3, p_4 \} = \{ p_1 * p_3, p_1 * p_4, p_2 * p_3, p_2 * p_4 \}$$

Рассматривая всякий путь p как $\{ p \}$, можно прийти к понятию исчисления путей графа, также имеющего структуру замкнутого полукольца $(M, *, +, \Theta, \{ \})^3$, где M – множество се- мейств путей между вершинами. Таким образом, можно сказать, что функция разметки (1) со

³ Не составляет труда проверить, что введенное исчисление удовлетворяет описанной ранее аксиоматике

свойствами (2) фактически осуществляет гомоморфизм полукольца семейств путей на полукольцо меток (см., например, [3]). И вернее (1) переписать следующим образом:

$$\lambda : M \rightarrow S .$$

Общая постановка задач о путях. Как отмечено в [2], большинство матричных алгоритмов на графах, использующих метод динамического программирования, имеют своей целью отыскание величин

$$c(i, j) = \bigoplus_{i \xrightarrow{P} j} \lambda(p) , \quad i, j \in V , \quad (6)$$

где “суммирование” проходит по всем путям из i в j , а $\lambda(p)$ есть метка пути p , т. е. “произведение” меток на ребрах.

Например, алгоритм вычисления кратчайших расстояний между парами вершин графа (алгоритм Флойда – Уоршола) предполагает использование в качестве меток *весов ребер*. Операции \otimes полукольца будет соответствовать суммирование весов ребер, а \oplus – минимизация весов альтернативных путей между соответствующими вершинами.

Широко известный алгоритм транзитивного замыкания графа использует в качестве множества $S \{0, 1\}$, операции \oplus – логическое ИЛИ, а \otimes – логическое И.

*Следовательно, каждый алгоритм предполагает задание соответствующего замкнутого полукольца меток $(S, \oplus, \otimes, \bar{0}, \bar{1})$ и гомоморфизма $\lambda(p)$ полукольца путей графа $(M, *, +, \Theta, \{\})$ на полукольцо меток.*

Пример применения алгебраического подхода к построению алгоритма. Поставим задачу нахождения последовательности ребер, составляющих оптимальный путь между конкретными двумя вершинами графа. Решим задачу путем расширения стандартного алгоритма Флойда – Уоршола.

Заметим прежде всего, что $\lambda(p)$ в указанном алгоритме является *эпиморфизмом* (т. е. *отображением на*), а значит, одной метке (весу) $s \in S$ может соответствовать несколько путей. Условие задачи предполагает выяснение состава только *оптимального пути*, следовательно, достаточно потребовать не полного изоморфизма полуколец меток и путей, а лишь моноидов $(M', *, \Theta)$ и $(S, \otimes, \bar{1})$, где $M' \subseteq M$ – подмножество одиночных путей в множестве семейств.

Введем вспомогательное определение

$L = \{l_1, l_2, \dots, l_n\} \subset M$ – множество образующих моноида $(A, *, 1)$, если $\forall a \in A$, $a = l_1^{f_1} * l_2^{f_2} * \dots * l_n^{f_n}$, где $f_1, f_2, \dots, f_n \in N$. Моноид, порожденный системой образующих L , обозначим как $\langle L \rangle$.

В частности, множеством образующих моноида $(M', *, \Theta)$ являются ребра графа. Очевидно, что в рамках решаемой проблемы гомоморфизм $\lambda(p)$ должен обеспечивать, прежде всего, однозначное отображение систем образующих рассмотренных ранее моноидов. Изоморфность $\lambda(p)$ будет, очевидно, обеспечена при совпадении мощностей множеств M' и S . Для графа с n ребрами этим условиям будет удовлетворять функция разметки, сопоставляющая каж-

дому ребру число из моноида, порожденного n взаимно простыми числами с помощью обыкновенного умножения. Аналогично, всякой конкатенации путей будет соответствовать произведение чисел из моноида меток, а такое число может быть однозначно разложено на сомножители. На основании данных соображений расширим используемое алгоритмом Уоршолла-Флойда полукольцо меток $(S', \oplus', \otimes', \bar{0}', \bar{1}')$ следующим образом:

Флойда полукольцо меток $(S', \oplus', \otimes', \bar{0}', \bar{1}')$ следующим образом:

1) Носитель полукольца $(S', \oplus', \otimes', \bar{0}', \bar{1}')$ S' определим как подмножество декартова произведения $S \times K$, где K – носитель моноида $(K, *, 1) = \langle L \rangle$. Таким образом, носитель будет множеством упорядоченных пар (k_i, s_i) – сложных меток, в которых первый компонент отражает состав пути, а второй – его стоимость.

2) Операция \otimes' будет сводиться к соответствующим операциям над компонентами сложных меток графа:

$$(k_1, s_1) \otimes' (k_2, s_2) = (k_1 * k_2, s_1 \otimes s_2).$$

3) Будем считать, что операция \oplus' полностью сводится к операции \oplus , т.е. выполняется лишь над вторыми компонентами сложных меток:

$$(k_1, s_1) \oplus' (k_2, s_2) = (k_1 | k_2, s_1 \oplus s_2),$$

где $k_1 | k_2$ обозначает выбор k_1 или k_2 в зависимости от результата $s_1 \oplus s_2$ (в данном случае выбор минимального из двух путей);

4) $\bar{0}'$ будет являться пара $(0, \bar{0})$;

5) $\bar{1}'$ будет пара $(1, \bar{1})$.

Реализация алгоритма основана на матрице Q , подобной матрице весов, используемой в алгоритме Уоршолла-Флойда:

$$Q = \left| (a_{ij}, s_{ij}) \right| \quad i = 1, \dots, n,$$

где n – число вершин.

Сам алгоритм сводится к получению ряда матриц $Q^2 = Q \times Q$, $Q^3 = Q^2 \times Q$, ..., где \times соответствует операции обычного умножения матриц с заменой умножения и сложения на соответствующие операции полукольца сложных меток. Элементы матрицы $Q^k = \left| (a_{ij}^k, s_{ij}^k) \right|$, полученной на k -й итерации, содержат сведения как о составе путей длины k ребер (компонент a_{ij}), так и о их стоимости (компонент s_{ij}).

Как и для большинства матричных алгоритмов, среднее время работы может быть оценено как $O(n^3)$, причем коэффициент при n^3 будет зависеть от выбора моноида $(K, *, 1)$. Для программной реализации удобнее заменить моноид, построенный на основе простых чисел, моноидом следующей структуры:

1) система образующих – $L = \{2^i\}, i = 1, \dots, n$, n – число вершин графа;

2) операция – побитовое ИЛИ;

3) в качестве единицы моноида выделено число 0, а элемента $\bar{0}$ полукольца сложных меток – элемент $(2^{n+1} - 1, \bar{0})$.

СПИСОК ЛИТЕРАТУРЫ

1. Ахо А., Хопкрофт Дж. Построение и анализ вычислительных алгоритмов. М.: Мир, 1979.
2. Кормен Т., Лейзерсон Ч., Ривест Р. Алгоритмы. Построение и анализ. М.: МЦНМО, 2001.
3. Курош А.Г. Лекции по общей алгебре. СПб.: Лань, 2005.
4. С. Ленг. Алгебра. М.: Мир, 1963.

N. V. Vasilyev, A. I. Yachin

USE OF THE ALGEBRAIC APPROACH FOR CONSTRUCTION OF ALGORITHMS ON COLUMNS

Opportunities of use of the algebraic approach to research and construction of algorithms on columns are considered.

Abstraction, algorithm, column, gomomorphism, the closed half ring, matrix algorithm, monoid, a way the column, epimorphizm

УДК 368:519

Ю. Д. Григорьев, Ле Динь Шон

О СТРАТЕГИЯХ УПРАВЛЕНИЯ РИСКАМИ В ЗАДАЧАХ ПЕРЕСТРАХОВАНИЯ

Дается краткий обзор стратегий управления рисками в задачах перестрахования. Рассмотрены различные задачи выбора эксцедентных стратегий перестрахования, зависящие от вида минимизируемой функции потерь, в том числе задачи минимизации суммы дисперсий и среднеквадратического отклонения рисков страховщика и перестраховщика, максимизации вероятности неразорения в моделях индивидуального и коллективного рисков.

Эксцедентное перестрахование, меры риска, модели индивидуального и коллективного риска, вероятность неразорения, характеристический коэффициент Лундберга, модель Крамера–Лундберга.

Выбор оптимальной стратегии управления в той или иной сфере деятельности является основной задачей при принятии решений. В статье обсуждается единый подход к управлению рисками, который встречается в страховании, теории обслуживания технических систем, теории запасов и т. д. Речь идет о так называемом пороговом, или эксцедентном, управлении, когда до определенного момента времени (до определенного уровня издержек, запасов и т. д.) осуществляется одна политика, а с момента превышения соответствующего порога – другая. В страховании риски регулируются франшизой и перестрахованием [3], в технической диагностике – периодическими проверками [1], [2], в теории запасов – оптимальной стратегией выбора точки заказа [9].

В работе [6] показана эквивалентность стратегий выбора оптимального периода периодических проверок и оптимального уровня условной франшизы и установлена аналогия между моделью безусловной франшизы и задачей «рождественской елки» в теории управления запасами. Все это свидетельствует о том, что эксцедентный тип управления различными стохастическими системами применяется весьма широко.

В данной статье ограничимся рассмотрением задач только из области страхования, объединяемых идеей эксцедентного управления по различным критериям оптимальности.

Моментные меры риска. Пусть задан исходный страховой риск X – неотрицательная случайная величина (с.в.) с функцией распределения (ф.р.) $F(x) = \mathbf{P}\{X \leq x\}$. Требуется так определить комонотонные риски страховщика $Y = g(X)$ и перестраховщика $Z = h(X)$, чтобы минимизировать некоторый функционал $\varphi(Y, Z)$ на множестве рисков $\{(Y, Z) : Y + Z = X\}$. Напомним, что риски называются комонотонными, если их функции дележа $g(\cdot)$ и $h(x) = x - g(x)$ являются одновременно монотонно убывающими или монотонно возрастающими. В теории страхования рассматривается второй случай.

Положим $g(x) = \min(x, r)$, $h(x) = \max(0, x - r)$. Следуя [5], определим следующие два функционала моментного типа:

$$\varphi_1(Y, Z) = \mathbf{D}Y + \mathbf{D}Z, \quad \varphi_2(Y, Z) = \mathbf{M}(Y - Z)^2.$$

Обозначим $\alpha = 1 - F(M)$, и пусть

$$A_r = \int_0^r [1 - F(x)] dx, \quad B_r = \int_r^\infty [1 - F(x)] dx, \quad P = \mathbf{M}X = A_r + B_r.$$

Теорема 1 [5]. Пусть $Y = \min(x, r)$, $Z = \max(0, x - r)$ и $\alpha = o(r^2)$, $r \rightarrow \infty$. Справедливы следующие утверждения: оптимальный уровень удержания $r_i = \arg \min_{0 \leq r < \infty} \varphi_i(r)$ является положительным корнем следующих уравнений:

$$i = 1: \alpha r = (1 - \alpha)(P - A_r) + \alpha A_r, \quad (1)$$

$$i = 2: \alpha r = (1 + \alpha)(P - A_r) - \alpha A_r, \quad (2)$$

Уравнение (1) получено в работе [10], уравнение (2) – в работе [5]. Отметим внешнее сходство этих уравнений. Для распределений Парето и Эрланга оптимальные стратегии r_1 согласно критерию $\varphi_1(Y, Z)$ исследованы в [4], для равномерного распределения по критериям $\varphi_1(Y, Z)$ и $\varphi_2(Y, Z)$ – в работе [5]. Во всех этих случаях функции потерь $\varphi_i(Y, Z)$ оказывались унимодальными, т. е. оптимальные стратегии r_i – единственными.

Вероятность неразорения. Важным показателем деятельности страховой компании, характеризующим ее меру риска, является вероятность неразорения, которая определяется по-разному в зависимости от рассматриваемой модели риска. Модель индивидуального риска определяется следующими соотношениями:

$$C(t) = cN, \quad X(t) = \begin{cases} 0, & t < t_0, \\ X_1 + X_2 + \dots + X_N, & t = t_0. \end{cases}$$

Это означает, что компания заключает N однотипных договоров страхования с одинаковым сроком действия t_0 и возможными выплатами по i -му договору X_i , премия за каждый договор равна c . Пусть $Z = cN - (X_1 + X_2 + \dots + X_N)$ – капитал компании в момент окончания срока действия договоров. Тогда

$$W_N = \begin{cases} 1, & t < t_0, \\ \mathbf{P}\{cN - (X_1 + X_2 + \dots + X_N) \geq 0\}, & t = t_0. \end{cases}$$

– вероятность неразорения.

Для вычисления W_N применяют приближенные методы. Основной из них – нормальная аппроксимация суммы $S = X_1 + X_2 + \dots + X_N$, согласно которой $(S - a)/\sigma \xrightarrow{d} \Phi(0, 1)$, где $a = \mathbf{MS} = \mathbf{NM}X_1$, $\sigma^2 = \mathbf{DS} = \mathbf{ND}X_1$. Тогда

$$W_N \cong \Phi\left(\frac{cN - a}{\sigma}\right).$$

Поскольку при эксцедентном перестраховании a и σ^2 зависят от r , то можно максимизировать W_N , решая задачу $(cN - a)/\sigma \rightarrow \max_{0 \leq M < \infty}$.

В модели коллективного риска капитал страховой компании в момент времени $t \geq 0$ равен

$$Z(t) = x + C(t) - \sum_{i=1}^{N(t)} X_i, \quad (3)$$

где $Z(t)$ – процесс риска, определяемый начальным капиталом компании $x \geq 0$, процессом поступления премий $C(t)$ и совокупным требованием выплат $X(t) = \sum_{i=1}^{N(t)} X_i$ на промежутке $[0, t)$. Наиболее часто в качестве $Z(t)$ рассматривают классический процесс риска, или процесс риска Крамера–Лундберга [11].

Определим момент разорения $T = \inf\{t > 0 : Z(t) < 0\}$, т. е. момент, когда впервые на интервале $[0, \infty)$ капитал $Z(t)$ становится отрицательным. Если $\sup_{0 \leq u \leq t} Z(u) > 0$ для всех $t \geq 0$, то полагаем $T = \infty$. Величина $\Psi(x, t) = \mathbf{P}\{T \leq t | Z(0) = x\}$ называется вероятностью разорения на конечном интервале времени $[0, t)$ при начальном капитале x , а величина

$\Psi(x) = \lim_{t \rightarrow \infty} \Psi(x, t) = \mathbf{P}\{T < \infty | Z(0) = x\}$ – вероятностью разорения на бесконечном интервале времени $[0, \infty)$ при начальном капитале x .

Пусть $m_1 = \mathbf{MX}$. При $c - \lambda m_1 > 0$ (условие чистой прибыли) и $B(0) = 0$ функция $W(x) = 1 - \Phi(x)$ удовлетворяет интегральному уравнению Крамера [11]

$$W(x) - \frac{\lambda}{c} \int_0^x K(x-s)W(s)ds = W(0),$$

с ядром типа свертки $K(t) = 1 - B(t)$ и постоянной правой частью $W(0) = 1 - \lambda m_1 / c > 0$.

В случае эксцедентного перестрахования вероятность неразорения W является функцией не только x , но и уровня r , т. е. теперь полагаем $W = W(r, x)$. Будем также обозначать через $W_r(x)$ вероятность неразорения как функцию от начального капитала x при фиксированном r , а через $W_x(r)$ – зависимость вероятности неразорения от r при фиксированном значении x . Очевидно, при этом предполагается равенство

$$W(r, x) := W_r(x) := W_x(r).$$

Пусть $W_r(0) = 1 - v_r A_r$, $v_r^{-1} = (1 + \theta)m_1 - (1 + \xi)B_r$, где θ, ξ ($\xi \geq \theta$) – нагрузки относительной безопасности страховщика и перестраховщика соответственно.

Теорема 2 [12]. Пусть $W_r(x)$ – вероятность неразорения страховщика в модели риска Крамера–Лундберга при эксцедентном перестраховании. Тогда $W_r(x)$ удовлетворяет интегральному уравнению

$$W_r(x) = v_r \int_{\max\{0, x-r\}}^x [1 - B(x-S)] W_r(S) ds + W_r(0). \quad (4)$$

Эффективность численного решения уравнения (4) методами квадратур, Монте-Карло и аппроксимации сплайнами для различных распределений риска исследована в [8]. Аналитическое исследование уравнения (4) в случае марковской модели (экспоненциальное распределение риска $B(y) = 1 - e^{-y}$) предпринято в [12].

Для записи решения $W_r(x)$ введем функции W_{kr} – финитные функции на $J_{kr} = [(k-1)r, kr)$ (функции, равные нулю вне J_{kr}). Функции $W_{kr}(x)$ назовем элементами $W_r(x)$. Для стратегии r в рассматриваемой модели перестрахования должна быть задана нижняя граница δ , обеспечивающая выполнение условия $0 \leq W_r(0) \leq 1$. Как нетрудно видеть, ее можно определить из уравнения связи $A_\delta/B_\delta = (\xi - \theta)/\theta$. Из условия $r \geq \delta$ также следует, что функция v_r является непрерывной, положительной и монотонно невозрастающей по r .

Теорема 3. Пусть задана модель риска Крамера–Лундберга (3) с фиксированным управлением $r \geq \delta$. Тогда для марковской модели Крамера–Лундберга (3) с функцией распределения размера требований $B(y) = 1 - e^{-y}$ элементы $W_{kr}(x)$ вероятности неразорения $W_r(x)$, $x \in J_{kr} = [(k-1)r, kr)$, $k = 1, 2, \dots$, допускают аналитическое представление, а именно:

а) если $v_r \neq 1$ и $a = e^r$, $b = v_r/(1-v_r)$, $c = a^{v_r} = e^{rv_r}$, то

$$W_{k+1,r}(x) = Q_r e^{-kr} \{ r_{k+1,1} + e^{(1-v_r)(kr-x)} \sum_{j=0}^k r_{k+1,j+2} (x-kr)^j \}, \quad k = 0, 1, \dots$$

$$\text{где } Q_r = \frac{W_r(0)}{1-v_r} = 1 - \frac{e^{-r}}{\theta - (1+\xi)e^{-r}}, \quad r_{11} = 1, \quad r_{12} = -v_r \text{ и (для } k \geq 2)$$

$$r_{kj} = \begin{cases} \frac{a^k - (-b)^k}{a+b}, & j=1, \\ (-1)^k v_r b^{k-1} + \alpha_k - \beta_k, & j=2, \\ \frac{v_r}{j-2} r_{k-1,j-1}, & 3 \leq j \leq k+1, \\ 0, & j \geq k+2. \end{cases}$$

$$\alpha_k = c \sum_{n=2}^k \left\{ \frac{r^{n-2}}{(n-2)!} \sum_{i=n}^k (-v_r)^{n-i} (i-2)! r_{k-1,i} \right\}, \quad \beta_k = \sum_{i=n}^k (-v_r)^{2-n} (n-2)! r_{k-1,n},$$

б) если $v_r = 1$ и $a = c = (1 + \xi)/\theta$, то

$$W_{k+1,r}(x) = Q_{k+1} \sum_{j=0}^{k+1} r_{k+1,j+1} (x - kr)^j, \quad k = 0, 1, \dots$$

$$\text{где } Q_{k+1} = \left(\frac{\theta}{1 + \xi} \right)^{k+1}, \quad r_{11} = 1, \quad r_{12} = 1, \quad r_{21} = a(r+1), \quad r_{22} = a-1, \quad r_{23} = -\frac{1}{2} \text{ и (для } k \geq 2)$$

$$r_{k+1,j} = \begin{cases} \frac{a(r+k)(-r)^{k-1}}{k!} + a^k \left[1 + kr + \sum_{i=1}^{k-2} (k-i) \left(1 + \frac{k-i}{i+1} r \right) \left(-\frac{r}{a} \right)^i \right], & j = 1, \\ a^k - r_{k,1}, & j = 2, \\ -\frac{1}{j-1} r_{k,j-1}, & 3 \leq j \leq k+2, \\ 0, & j \geq k+3. \end{cases}$$

Численное исследование показывает, что $W_x(r)$ является унимодальной функцией r . Таким образом, существует единственная оптимальная стратегия эксцедентного типа, максимизирующая вероятность неразорения $W(r, x)$ при заданном начальном капитале x . Отметим, что максимум $W_x(r)$ по r в зависимости от x может быть как гладким, так и негладким.

Перечисленные способы выбора эксцедентных стратегий перестрахования не исчерпывают всего многообразия вариантов оптимизации, возникающих на практике. В статье сделана попытка рассказать об эксцедентном подходе к проблеме оптимизации риска, при этом не были рассмотрены эксцедентные схемы перестрахования с двумя и более точками удержаний, вопросы упорядочения рисков страховщика и перестраховщика и многие другие вопросы.

Эти и другие вопросы пропорционального и непропорционального перестрахования в настоящее время активно обсуждаются в многочисленных журнальных и интернетовских публикациях, что свидетельствует о важности вопроса, затронутого в данной статье.

СПИСОК ЛИТЕРАТУРЫ

1. Барзилович Е. Ю. Модели технического обслуживания сложных систем. Учеб. пособие. М.: Высш. шк., 1982.
2. Барлоу Р., Прошан Ф. Математическая теория надежности. М.: Сов. радио, 1969.
3. Голубин А. Ю. Математические модели в теории страхования: построение и оптимизация. М.: АНКЛ, 2003.
4. Григорьев Ю. Д. Две теоремы об оптимальном уровне удержания для моделей перестрахования Борча и Штрауба // Тр. III Всерос. ФАМ конф. Красноярск: Гротеск, 2004. С. 94-100.
5. Григорьев Ю. Д. Дележи и оптимизация страхового риска // Тр. IV Всерос. ФАМ конф. Красноярск: Гротеск, 2005. С. 69-76.
6. Григорьев Ю. Д., Ле Динь Шон. Об одном общем подходе к управлению стохастическими системами страхования, технического обслуживания и управления запасами // Информационные технологии и системы (управление, экономика, транспорт): Межвуз. сб. науч. тр. Вып. 1. СПб.: ООО «Андреевский издательский дом», 2005. С. 41-48.
7. Ле Динь Шон. Выбор оптимального уровня удержания в схеме эксцедентного перестрахования // Тр. IV Всерос. ФАМ конф. – Красноярск: Гротеск. 2005. С. 98-103.
8. Ле Динь Шон. Сравнение различных численных методов решения интегрального уравнения Крамера в задаче эксцедентного перестрахования // Информационные технологии и системы (управление, экономика, транспорт): Межвуз. сб. науч. тр. Вып. 1. СПб.: ООО «Андреевский издательский дом», 2005. С. 76-83.

9. Рыжиков Ю. И. Теория очередей и управление запасами. СПб.: Питер, 2001.
10. Borch K. An Attempt to Determine the Optimum Amount of Stop Loss Reinsurance. Transactions of the XVI International Congress of Actuaries, 1960. Vol.2, P. 597-610.
11. Bühlman H. Mathematical Models in Risk Theory. Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Band 72, Springer Verlag: Berlin Heidelberg New York, Second Printed 1996.
12. Grigor'ev Yu. D., Le Dinh Son. Numerical Optimization of Value-at-Risk Functional for Stop Loss Reinsurance in the Cramer-Lundberg Risk Model // Modeling and Analysis of Safety and Risk in Complex Systems: Proc of the V Intern Scientific School MA SR, 2005, Saint-Petersburg, Russia, June 28 - July 1, 2005. P. 216-222.

Y. D. Grigor'ev, Le Dinh Son

ON THE STRATEGIES OF RISK MANAGEMENT IN THE REINSURANCE PROBLEMS

The review of the strategy of risk management in the reinsurance problems is presented. The various problems of an excess of loss reinsurance strategies depending on different loss functions are considered. As a case of point the sum of the risk variations and the risk mean square deviation of the insurer and the reinsurer, non-ruin probability in the individual and Cramer-Lundberg collective risk models are discussed.

Excess of loss reinsurance, risk measures, the individual and collective risk models, non-ruin probability, Lundberg's adjustment coefficient, the Cramer-Lundberg's model.

УДК 37.014.1

Мотасем АбуДавас, Н.Е.Новакова

ИНТЕЛЛЕКТУАЛЬНАЯ ПОДДЕРЖКА ПРОЕКТНЫХ РЕШЕНИЙ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ ХРАНИЛИЩ ДАННЫХ И DATA MINING

Рассматривается процесс аналитической обработки проектных данных. Предложена модель обучения выборки данных на основе генетических алгоритмов и нейронных сетей.

Знания, управление знаниями, интеллектуальные системы, проектная деятельность, базы знаний, раскопка и обнаружение знаний, хранилища данных, бизнес-анализ

В последние годы наблюдается тенденция все более тесной интеграции систем автоматизированного проектирования с другими автоматизированными системами и производственной сферой [1]. На смену идеям гибких производственных систем и компьютеризированного интегрированного производства пришла концепция CALS-технологий (Continuous Acquisition and Life cycle Support) или ИПИ (Информационная Поддержка жизненного цикла Изделий) [2]. Анализ развития информационных технологий в производственных задачах показывает, что одним из направлений движения является *все более полный охват всех стадий жизненного цикла продукции*. Жизненный цикл (ЖЦ) промышленных изделий включает ряд этапов, начиная от зарождения идеи нового продукта до утилизации по окончании срока его использования. К основным этапам ЖЦ промышленной продукции относятся анализ рынка, проектирование, технологическая подготовка производства, собственно производство, реализация продукции, эксплуатация и, наконец, утилизация (рис. 1).

Одна из задач CALS-технологий – унификация и стандартизация спецификаций промышленной продукции на всех этапах ее ЖЦ. Основные спецификации изделия представляются проектной, конструкторской, технологической, производственной, маркетинговой и эксплуатационной документацией. Применение CALS-технологий дает возможность предприятиям су-

шественно сократить объемы проектных работ, так как описания многих составных частей оборудования, машин, систем, радиоэлектронных схем, проектировавшихся ранее, должны храниться в унифицированных форматах данных. Организация полного ЖЦ изделия невозможна без *интегрированной информационной системы (ИИС), поддерживающей весь жизненный цикл изделия*. Основными компонентами ИИС являются хранилища данных, построенные на основе баз данных (БД) оперативной обработки информации и систем бизнес-аналитики.



Рис. 1

Для реализации основных концепций CALS и обеспечения полного ЖЦ изделия необходимо соблюдение основных принципов CALS-технологий: анализа и реинжиниринга бизнес-процессов, представления данных об изделии в электронном виде и интеграции данных об изделии. В результате применения основных концепций CALS можно получить огромное число вариантов конфигураций системы. Для оценки технико-экономических характеристик таких автоматизированных подсистем как САПР и АСУ ТП, а также эффективности организации бизнес-процессов полного ЖЦ изделия необходимы соответствующие инструментальные средства. В статье предлагается система поддержки проектных решений (СППР), которая может быть полезна разработчикам подобных систем на различных этапах ее ЖЦ. Архитектура СППР показана на рис. 2. Цифрами обозначены следующие подсистемы: 1 – обработка экспертных оценок; 2 – оценка системы на основе статистики прошлых проектов; 3 – имитационное моделирование; 4 – оценка сложности проекта; 5 – БД оценок экспертов; 6 – БД статистики проектов; 7 – БД моделей компонентов. В системе используются различные методы моделирования и анализа бизнес-процессов. В качестве инструментальных средств обработки результатов моделирования применяются: метод экспертных оце-

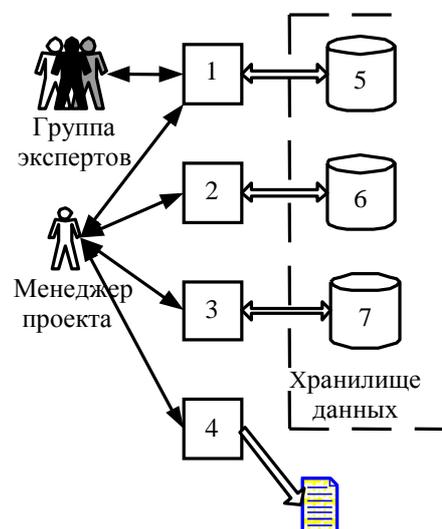


Рис. 2

нок. В качестве инструментальных средств обработки результатов моделирования применяются: метод экспертных оце-

нок, аналитические средства обработки больших объемов проектных данных на основе OLAP-решений и технологий Data Mining, а также имитационное моделирование. Для реализации каждого из перечисленных методов применяется своя БД, обеспечивающая информационную поддержку процесса моделирования. Обозначенные на рис. 2 цифрами 5, 6 и 7 базы данных объединены в общее хранилище данных, обеспечивающее единое информационное пространство для организации процесса поддержки проектных решений.

Используемые в СППР различные методы моделирования позволяют изучить проблему с разных точек зрения. *Метод экспертных оценок* дает наименее достоверные и наиболее субъективные результаты. Его основное преимущество – невысокая стоимость реализации. Использование технологий *интеллектуального анализа данных* обеспечивает обработку больших объемов информации, поддерживая стандартную форматированную отчетность и различные аналитические инструменты – OLAP, выявление скрытых данных, планирование и моделирование. Основным недостатком этого подхода заключается в использовании больших объемов статистических данных. Это особенно важно для обработки данных с использованием алгоритмов Data Mining. Технологии интеллектуального анализа данных особенно эффективны для предприятий, в которых организовано хранение архивных данных по проектам прошлых лет. Имитационное моделирование представляет мощный гибкий механизм изучения процессов различной природы. Имитационный эксперимент позволяет специалистам, исследующим бизнес-процессы, за относительно небольшой промежуток времени просмотреть множество вариантов архитектуры системы и выбрать наиболее оптимальный по заданной системе критериев. Перечисленный список методов моделирования может быть расширен. Система построена таким образом, что в случае необходимости можно добавлять новые виды моделирования и анализа конфигурации САПР. Данная статья посвящена применению методов интеллектуального анализа данных.

В соответствии с концепциями CALS-технологий [1] особую актуальность приобретают вопросы комплексирования средств автоматизации проектирования (АП) в единую систему и обеспечение интеграции САПР и других систем автоматизации жизненного цикла изделия. Задача внедрения САПР или ее модернизации требует значительных усилий коллектива высококвалифицированных специалистов. Ошибки в принятии проектных решений, допущенные на ранних стадиях создания системы, особенно в проектировании архитектуры, могут привести к неудачам в проекте, а, возможно, и к полному его провалу. Одним из эффективных средств организации информации на основе технологии хранилищ данных является создание информационно-аналитических систем интеллектуальной поддержки проектных решений [3]. Организация информации в хранилище осуществляется с использованием *интегрированных платформ* (виртуального пространства) управления знаниями. Единый доступ к разнородным данным реализуется через модуль Knowledge Management, обеспечивающий управление, классификацию и представление общего набора сервисов для хранилищ информации [4].

Центральным компонентом *виртуального пространства* в системах интеграции информации является инструментарий интеллектуального анализа данных (Business Intelligence – BI). С его помощью можно создавать хранилища данных и управлять ими, обеспечивая обработку больших объемов информации, поддерживая стандартную форматированную отчетность и различные аналитические инструменты – OLAP, выявление скрытых данных, планирование и моделирование (рис. 3). При этом для каждого потре-

бителя информации (проектировщика, конструктора, менеджера проекта, аналитика, системного администратора) организуется оптимизированная рабочая среда аналитической обработки данных и накопленных знаний [5].

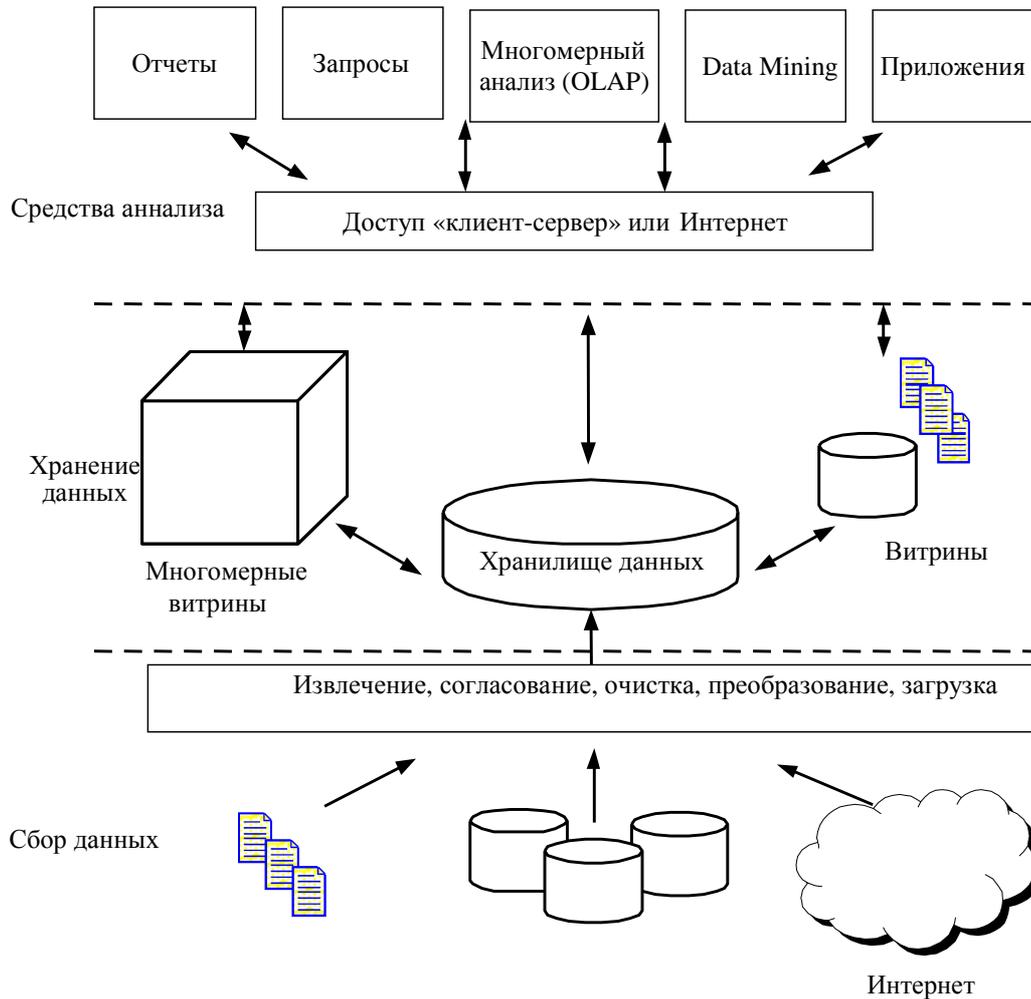


Рис. 3

Реализация системы интеграции информации требует согласованности данных в имеющихся хранилищах, особенно в географически распределенных компаниях и подразделениях контрагентов. Организация инфраструктуры технологии интеграции неоднородной информации предусматривает исключение неизбежного в этом случае несоответствия и дублирования. На пути решения этой проблемы разрабатываются информационные технологии, обеспечивающие свободную конфигурацию и способные быстро адаптироваться к изменяющимся условиям [2].

При разработке блока аналитических решений для задач интеллектуальной поддержки проектировщика большое значение имеет процесс управления информацией. Автоматизация этого процесса предусматривает участие проектировщика, особенно в ситуациях, когда оперативной информации недостаточно и необходимо обращаться к хранилищам данных. Большой объем информации требует мощных поисковых систем, способных обеспечить доступ к разнородным источникам данных, а также эффективных средств их визуализации.

В интеллектуальную среду поступает первичный материал – оперативные данные из различных систем (рис. 4). При дальнейшей обработке данных происходит ряд преобразований: данные → информация → знания → планы → действия.



Рис. 4

Хранилище данных извлекает данные из множества транзакционных или оперативных систем, а затем интегрирует и сохраняет данные в специализированной реляционной базе или в файле. Такой процесс извлечения и интеграции преобразует данные в новый продукт – информацию, т. е. выполняется первый шаг цикла «от данных к информации».

Затем пользователи, работающие с аналитическими инструментами (например, для создания запросов, отчетов, OLAP-анализа и выполнения операций data mining), обращаются к информации из хранилища и анализируют ее. Таким образом, удастся выявить тенденции, структуры и исключения. Аналитические инструменты помогают пользователям преобразовать *информацию в знания*.

Вооруженные такими познаниями пользователи могут создавать правила на основе обнаруженных тенденций и структур. Правила могут быть простыми (например, «Заказать 50 новых единиц, если на складе осталось менее 25») в виде прогнозы или сценарии «что если», опирающихся на тенденции и рабочие предположения, а также очень сложными, основанными на статистических алгоритмах или моделях. Например, статистически-сгенерированные правила могут динамически конфигурировать цены в ответ на изменившиеся условия рынка, оптимизировать расписания транспортировки грузов в крупной сети перевозок, или определять наиболее благоприятные условия перекрестных продаж, используя модели отклика клиентов. Таким образом, происходит переход от *знаний к правилам*.

На основе выявленных правил пользователи могут создавать планы и делать прогнозы. Например, специалисты по маркетингу разрабатывают специальные кампании, сформированные на основе анализа потребительских сегментов, моделей, прогнозирующих реакцию клиентов на конкретные предложения, и результатов предыдущих кампаний. В этих кампаниях указывается, какие предложения делать конкретным клиентам по различным каналам (например, через почтовую или электронную рассылку). Затем при выполнении плана знания и правила преобразуются в действия, и происходит преобразование «от правил к планам и действиям».

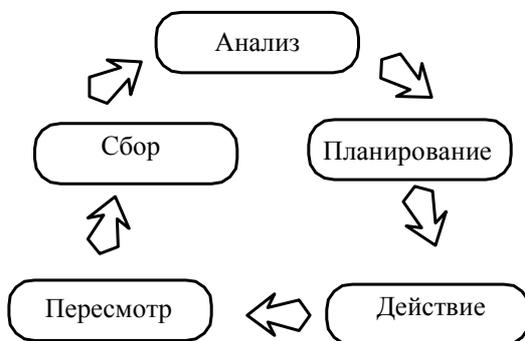


Рис. 5

После выполнения плана цикл повторяется. Оперативные системы собирают информацию о реакции клиентов на предложения или план, а также данные о последующих транзакциях (например, о продажах). Затем эти данные извлекаются хранилищем, интегрируются с соответствующей информацией из других источников и анализируются пользователями, которые оценивают эффективность своих планов и дорабатывают их нужным образом. Затем цикл снова повторяется.

Наибольшего эффекта при использовании CALS-технологий можно достичь при постоянном «обучении» компании, которая способна гибко и легко реагировать на любые изменения на рынке, состоящем из пяти этапов (рис. 5). Такой цикл создает так называемый

мую обучающуюся компанию. Когда организация повторяет этот цикл обучения, у руководства и сотрудников возникает четкое понимание того, как работает бизнес, как их решения и действия влияют на рынок, и наоборот.

Лучшие BI-решения обеспечивают мощную поддержку каждого этапа цикла обучения, дают организации возможность предвидеть изменения на рынке.

Традиционная математическая статистика, долгое время претендовавшая на роль основного инструмента анализа данных, откровенно спасовала перед лицом возникших проблем. Главная причина – концепция усреднения по выборке, приводящая к операциям над фиктивными величинами. Методы математической статистики оказались полезными главным образом для проверки заранее сформулированных гипотез (verification-driven data mining) и для «грубого» разведочного анализа, составляющего основу оперативной аналитической обработки данных (online analytical processing, OLAP). Основные аналитические инструменты, удовлетворяющие перечисленным требованиям, сегодня относят к области технологий Data Mining (раскопки данных).

Основу системы поддержки принятия проектных решений составляет среда хранилища данных (рис. 6), представленная в виде двух пересекающихся областей. Область, расположенная слева, называется «средой хранилища». Это тот участок работы, на который технические специалисты тратят от 60 до 80 % времени. Их задача состоит в извлечении, очистке, моделировании, преобразовании, передаче и загрузке транзакционных данных из одной или нескольких оперативных систем в хранилище.

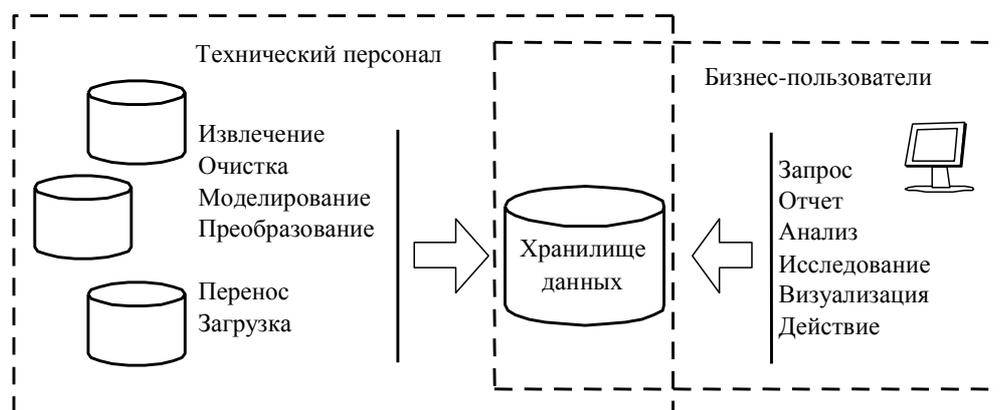


Рис. 6

Реализация этих функций в хранилище довольно сложная, так как операционные данные редко бывают безошибочными и согласованными, а также плохо поддаются интеграции. Техническим специалистам приходится расшифровывать смысл и значение тысяч элементов данных и величин во множестве операционных систем. Затем необходимо все интегрировать в единую согласованную модель бизнеса. Завершив очистку и анализ данных, технический персонал загружает интегрированные данные в хранилище, которое обычно представляет собой реляционную базу, оптимизированную для обработки запросов и генерации отчетов [5]. Часто для нужд пользователей конкретного подразделения приходится создавать специализированное подмножество хранилища, которое называют витриной данных. Витрину можно реализовать в виде реляционной базы или специализированной многомерной базы, которая позволяет пользователям выполнять выборку данных с заданными значениями измерений (slice and dice). Специалистам по маркетингу могут потребоваться такие измерения, как клиент, география, время, доход и т. п. Специали-

стов по разработке и внедрению САПР интересуют такие характеристики, как загрузка оборудования, среднее время проектирования изделия, типовой маршрут проектирования. Проектировщикам и конструкторам, выполняющим проектную задачу, необходима специализированная информация, характеризующая конкретную проектную ситуацию.

Правая часть на рис. 6 описывает аналитическую среду. Основной контингент специалистов, работающих с аналитикой, – бизнес-пользователи, которые применяют инструменты для выполнения запросов, отчетов, а также анализа, исследования, визуализации и (что самое главное) действуют, опираясь на данные в хранилище. Так как большинство бизнес-пользователей работают со стандартными отчетами, то отчеты создаются заранее техническим персоналом и размещаются в корпоративной сети. Отчеты можно просматривать в виде статических документов, устанавливать фильтры по соответствующим критериям или перемещаться в разных направлениях (выполнять поиск, углубление в данные, переход от одного измерения к другому) с тем, чтобы изменить представление или уровень детализации. До недавнего времени BI-системы собирали транзакционные данные, периодически делая «снимки» данных в учетной системе в определенное время дня или недели. Однако сейчас компании стараются добиться более эффективного принятия оперативных решений (в отличие от стратегических и тактических решений, которые обсуждаются далее), анализируя интегрированные данные более своевременно. Например, проектировщикам нужно рассмотреть результаты за вчерашний день, а не за прошлый месяц, сравнить эффективность работы в тот же день в прошлом году по аналогичным деталям узлов или механизмов, входящим в состав ранее спроектированных изделий, и т. п.

Для поддержки такого оперативного принятия решений BI-системы постепенно приближаются по своим характеристикам к транзакционным. Команды разработчиков используют «активные Хранилища», «оперативные склады данных» и промежуточное программное обеспечение (в том числе применяя интеграцию прикладных систем предприятия – EAI, Enterprise Application Integration – и Web-сервисы) для сбора данных практически в реальном времени и максимально быстро передают эти данные бизнес-пользователям. В этом случае полезно предоставить пользователям «бизнес-панели управления», которые работают в режиме, близком к реальному времени. Это даст возможность пользователям наблюдать состояние процесса или события, наблюдая за изменениями на счетчиках и шкалах панели.

В этой связи в рассматриваемой области в последнее время стали развиваться методы «неклассического» интеллектуального анализа данных, получившие название Data Mining и Knowledge Discovery (DM&KD). Как отмечается в [2], DM&KD представляет собой процесс управляемого данными (data-driven) извлечения зависимостей из больших баз данных. В этом процессе центральное место занимает порождение характеризующих анализируемые данные моделей, правил или функциональных зависимостей, которые затем представляются пользователю для оценки их релевантности, полезности и практической ценности. При этом под «зависимостями» понимаются не только традиционные аналитические соотношения, но и выделение кластеров (групп) или каких-либо шаблонов в данных, построение классификаций и порождение ассоциаций, установление причинных отношений и т. д. При этом применяются методы порождения деревьев решений, рассуждения на основе правил, генетические алгоритмы, нейронные сети, когнитивная графика. Как следствие, в существующих системах DM&KD пользователю предлагаются приближенные, «не точные» и «не оптимальные» решения, кото-

рые, тем не менее, позволяют получать важные результаты, правила и обобщения. Разведка (раскопка) данных – это процесс обнаружения в сырых данных ранее неизвестных, нетривиальных, практически полезных, доступных для интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

После извлечения из сырых данных зависимых параметров для некоторых целевых характеристик часто требуется уменьшить количество зависимых значений. Во многих ситуациях неизвестно наверняка, какие параметры, извлекаемые из данных, влияют на решение проблемы. Использование таких параметров увеличивает как время работы многих алгоритмов, так и время на выборку зависимых параметров. Так как количество примеров, необходимых для получения шаблона, пропорционально размерности набора параметров, уменьшение числа параметров будет также способствовать уменьшению количества необходимых примеров. Параметры могут по-разному влиять на решение, и им должны быть присвоены соответствующие веса при решении задач Data Mining. Следовательно, при поиске определенных шаблонов требуется идентификация и выбор «полезных» наборов параметров.

Можно выделить 2 типа задач, которые с разной эффективностью решаются разными методами Data Mining (хотя реальные задачи исследования данных могут совмещать в себе оба типа): описательные и прогнозирующие (предсказательные). Оба типа задач могут быть полезны на разных этапах сквозного цикла проектирования промышленных изделий. Нахождение, построение из данных различных моделей, которые могли бы быть использованы для прогноза и принятия решения в будущем, при встрече ситуации, отсутствовавшей в данных при выводе модели. К наиболее типичным задачам Data Mining относятся классификация, кластеризация, ассоциация, регрессия. Для решения задач Data Mining используются различные методы: алгоритмы, основанные на переборе, нечеткая логика, генетические алгоритмы [3], [6]. Применение генетических алгоритмов (ГА) при решении задач раскопки данных состоит в комбинации выборки набора параметров с неким обучаемым алгоритмом. При таком подходе оптимальность набора параметров, получаемого в процессе эволюционного вычисления, оценивается обучаемым алгоритмом. Существуют различные подходы к кодированию индивида популяции. Простейший вариант – представление наличия параметра в наборе в двоичном коде: 1 – параметр в наборе, 0 – параметр не в наборе.

Позже появились расширенные варианты представления индивидов – сериями весов от 0 до 10, в зависимости от важности параметра. Такой подход дает лучшие результаты, чем двоичное кодирование, на реальных, зашумленных наборах данных.

Для реализации селекции могут использоваться такие алгоритмы, как турнирная или рулеточная селекция, селекция по ранжиру, стабильная селекция.

При использовании турнирной селекции каждый турнир строится на выборке k элементов из популяции и выбора лучшей особи среди них. Наиболее распространен турнирный отбор с $k = 2$

Алгоритм работы оператора турнирной селекции включает следующие шаги:

1) Из популяции случайным образом выбирается k уникальных индивидов для участия в турнире: $e_1 \dots e_k$.

2) Из k индивидов выбирается один e_r с наибольшим значением фитнеса $f(e_r) = \max(f(e_1) \dots f(e_k))$.

3) Полученный индивид e_r считаем выбранным для участия в операции кроссовера.

При использовании рулеточной селекции вначале каждому индивиду назначается вероятность выбора, равная отношению его приспособленности (фитнеса) к суммарной приспособленности популяции. Рулетка отбирает особей с помощью «запусков» рулетки. Колесо рулетки содержит по одному сектору для каждого члена популяции. Размер каждого сектора пропорционален соответствующей величине фитнеса индивида. При таком отборе члены популяции с более высокой приспособленностью с большей вероятностью будут чаще выбираться, чем особи с низкой приспособленностью.

Алгоритм работы оператора рулеточной селекции:

- 1) Вычисляется сумма фитнесов всех индивидов популяции $S = \sum f(e)$.
- 2) Для каждого индивида вычисляется значение вероятности его выбора $p_e = f(e)/S$.
- 3) Популяция сортируется по значению p .
- 4) Вычисляется случайное значение r из диапазона $[0, S]$.
- 5) Осуществляется обход индивидов популяции, начиная с тех, что имеют минимальный фитнес, $sum = 0$.
- 6) Суммируются значения фитнеса, $sum = sum + f(e)$.
- 7) Если $sum < r$, то переход на итерацию 6.
- 8) Полученный индивид считаем выбранным для участия в операции кроссовера.

Рулеточный отбор не подходит для популяций индивидов с большим разбросом значений фитнес-функции. Например, если лучший индивид покрывает 90 % от колеса рулетки, вероятность выбора других индивидов для селекции крайне мала. Селекция по ранжиру присваивает определенное значение каждому индивиду популяции. Это значение пропорционально значению фитнес-функции рассматриваемого индивида. Чем больше значение фитнес-функции, тем большее значение ему присваивается (худшему индивиду – значение 1, лучшему – N , где N – количество индивидов в популяции). Таким образом, все гены популяции имеют вероятность быть выбранными.

Алгоритм работы оператора селекции по ранжиру:

- 1) Популяция сортируется по значению $f(e)$.
- 2) Каждому индивиду присваивается значение p из диапазона $[1, N]$, где N – численность популяции. Чем больше у индивида фитнес, тем большее значение p ему присваивается.
- 3) Вычисляется случайное значение r из диапазона $[1, N]$.
- 4) Осуществляется обход индивидов популяции, начиная с тех, у которых значение p минимально.
- 5) Суммируются значения p , $sum = sum + p$.
- 6) Если $sum < r$, то переход на итерацию 5.
- 7) Полученный индивид считаем выбранным для участия в операции кроссовера.

В каждой популяции выбираются лучшие индивиды, из них создаются новые индивиды. Наихудшие индивиды замещаются созданными, а остальные помещаются в новую популяцию без изменений.

Алгоритм работы оператора стабильной селекции:

- 1) Популяция разбивается на 2 подмножества: E – подмножество индивидов с большим фитнесом, предназначенная для создания потомства и B – подмножество индивидов для замены.
- 2) Выбирается индивид e_b из подмножества B .
- 3) Из подмножества E случайным образом выбираются родители e_{p1} и e_{p2} .

4) Над элементами e_{p1} и e_{p2} проводится операция кроссовера, в результате которой получается потомок e_{ch} .

5) Элемент e_b заменяется элементом e_{ch} .

6) Если в подмножестве B остались необработанные элементы, то переход на итерацию 2.

Применение корпоративного хранилища данных как основы интеллектуальной поддержки проектных решений позволит повысить эффективность автоматизированного проектирования промышленных изделий. Кроме того, хранилище данных является информационной основой для применения CALS-технологий и организации виртуальных производств.

СПИСОК ЛИТЕРАТУРЫ

1. Норенков И. П. Основы автоматизированного проектирования: Учеб. для вузов, 2-е изд., перераб. и доп. М.: Изд-во МГТУ им. Н. Э. Баумана, 2002.
2. Новакова Н. Е. Языки функциональных спецификаций для предметных областей САПР: Учеб. пособие. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2005.
3. А. А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод. Методы и модели анализа данных: OLAP и Data Mining. СПб.: БХВ-Петербург, 2004.
4. Горячев А. В. Новакова Н. Е. Управление знаниями в проектной деятельности // Изв. СПбГЭТУ «ЛЭТИ», 2004, Вып. 1, с. 29-39.
5. Спирли Э. Корпоративные хранилища данных. Планирование, разработка, реализация, пер. с англ. М.: Изд. дом «Вильямс», 2001.
6. Дюк В., Самойленко А. Data Mining. СПб.: Питер, 2001.

Motasem AbuDavas, N. E. Novakova

INTELLECTUAL SUPPORT OF THE DESIGN DECISIONS WITH IMPLEMENTATION DATA WAREHOUSE AND DATA MINING TECHNOLOGIES

By development of information-analytical system for intellectual support of the design decisions on the basis of technology of the Warehouses, the important meaning has revealing both analysis of the «latent» data and knowledge on the basis of the concept Data Mining.

Knowledge, knowledge management, knowledge-based systems, project activity, knowledge base, Data Mining, knowledge discovery, data warehouse, business intelligence

УДК 681.3

О. И. Кутузов, Т. М. Татарникова

ПОДХОД К ОПТИМИЗАЦИИ СТРУКТУРЫ МЕЖСЕТЕВОГО УСТРОЙСТВА С ПРИВЛЕЧЕНИЕМ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

В рамках решения проблемы организации обмена данными между сетевыми объектами разработана концепция настраиваемой модульной структуры межсетевого устройства. Предлагаемая модульная технология обеспечивает направленное формирование облика межсетевого устройства с заданным набором свойств и позволяет расширять его конфигурацию. Основу предлагаемой процедуры составляет специально разработанный генетический алгоритм оптимизации структуры межсетевого устройства по производительности.

Межсетевое устройство, мультипроцессорная система, параллельная обработка, оптимизация структуры устройства, генетический алгоритм

Межсетевые устройства (МУ), такие как, центры коммутации пакетов, концентраторы, маршрутизаторы, шлюзы, серверы, число которых в распределенной системе может достигать несколько сотен, выполняются в виде многопроцессорных вычислительных сис-

тем (МВС), обеспечивающих необходимую производительность за счет параллельной работы ее элементов и оптимального распределения нагрузки между модулями МВС [1].

Одним из преимуществ параллельных вычислительных комплексов является их способность адаптироваться к внешним условиям, в частности, за счет динамического перераспределения вычислительной мощности в зависимости от приоритета решаемых задач.

Новая технология построения высокопроизводительных МУ, описываемая в данной статье, основана на модульном принципе построения МВС, что позволяет пользователям расширять свои системы по мере необходимости. Система, построенная на базе такой технологии, включает «кирпичики» различных типов:

процессорный модуль, содержащий процессоры обработки пакетов (ПОП) вместе с локальной памятью для хранения соответствующих программ (протоколов);

общее поле памяти (ОПП), разбитое на секции (СОП) для обеспечения параллельного доступа нескольких ПОП к разным СОП;

модуль базового ввода-вывода, представляющий собой набор из Y процессоров портов (ПП) межсетевого устройства;

коммутационный модуль, который содержит центральный блок комплексирования всех модулей устройства между собой;

блок расширения, содержащий слоты горячей замены.

Новые типы модулей могут быть добавлены в систему для построения специализированных конфигураций.

Каждый ПОП выполняет строго определенные протоколы: ПОП 2 реализует протокол 2-го уровня, ПОП 3 – протокол 3-го уровня и т. д. до 7-го уровня эталонной модели взаимодействия открытых систем. Здесь возможна параллельная обработка пакетов на разных ПОП как одного и того же уровня, так и разных уровней. При этом количество ПОП разных уровней не обязательно должно быть одинаковым. Например, если на выполнение процессов в ПОП $(n + 1)$ -го уровня необходимо больше времени, чем в ПОП n -го уровня, то для исключения неограниченного роста очереди к ПОП $(n + 1)$ в работу включается еще один (или несколько) ПОП $(n + 1)$ -го уровня.

Каждый пакет, записанный процессором порта в ОПП, генерирует задание, которое выполняется на ПОП по некоторому алгоритму, определяющему последовательность обработки пакетов процессами [2].

В МУ в реальном масштабе времени выполняется ряд процессов, реализующих заданные функциональные алгоритмы. Каждый пакет, записанный ПП в ОПП, генерирует задание, которое задается управляющей информацией, содержащейся в заголовке пакета, и которое обрабатывается прикладными процессами по некоторому последовательному алгоритму (технологической цепочке). Информационную связь процессов, выполняемых на ПОП, отражает рис. 1.

Одновременно, если позволяет число ПОП, может обрабатываться несколько заданий (пакетов). Если задание, приписанное некоторому прикладному процессу, не находит свободного ПОП, то оно ставится в очередь заданий. По освобождении необходимого ПОП он приписывается первому ожидающему его заданию.

Передача задания от одного процесса к другому, распределение активных процессов по ПОП, регулирование доступа процессов к ресурсам осуществляется операционной системой. Последняя состоит из набора макрокоманд, позволяющих выполнять все необходимые систем-

ные операции. Обращение прикладного процесса к любой макрокоманде инициирует на ПОП обращение к ресурсам, находящимся в ОПП. Если при запросе некоторого ресурса он оказывается занятым, то процесс ставится в очередь к данному ресурсу и его дальнейшее выполнение блокируется. ПОП, на котором выполнялась обработка пакета, освобождается. При освобождении ресурса он приписывается первому ожидающему его заблокированному процессу, и обработка пакета продолжается с того места алгоритма обработки, где она была прервана.

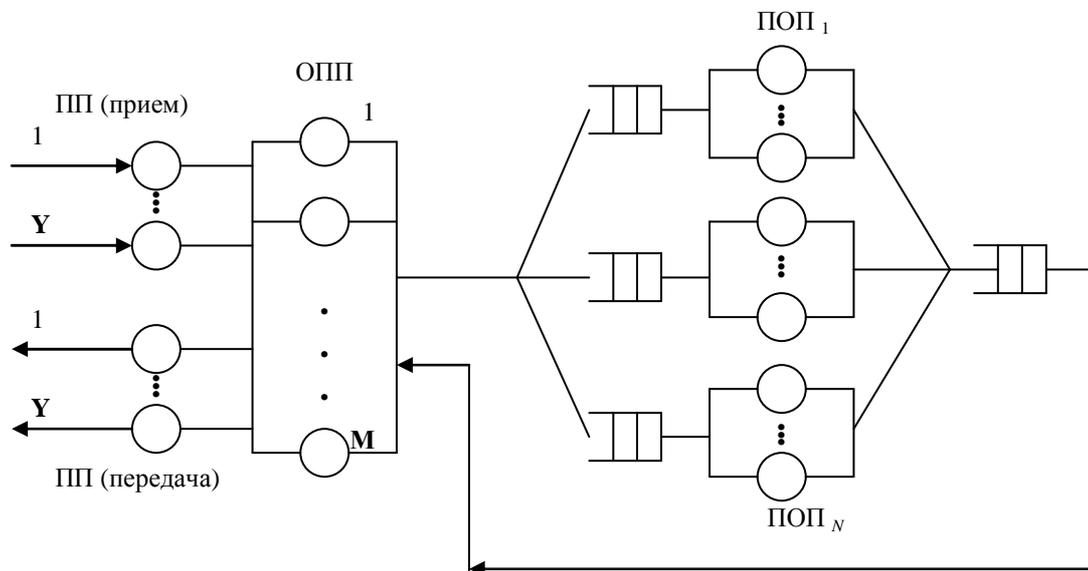


Рис. 1

Для определения оптимальной структуры межсетевое устройства предлагается привлечь генетические алгоритмы, которые являются очень популярным в настоящее время способом решения задач оптимизации. В их основе лежит использование эволюционных принципов для поиска наилучшего решения [3].

Рассмотрим работу генетического алгоритма. Представим себе искусственный мир, населенный множеством существ (особей), причем каждое существо – это некоторое решение поставленной задачи. Будем считать особь тем более приспособленной, чем лучше соответствующее решение (чем лучшее значение целевой функции оно дает). Тогда задача оптимизации сводится к поиску наиболее приспособленного существа. Конечно, невозможно поселить в наш виртуальный мир все существа сразу, так как их очень много. Вместо этого рассматривается много поколений, сменяющих друг друга. Теперь, если ввести в действие естественный отбор и генетическое наследование, то полученный мир будет подчиняться законам эволюции. Заметим, что в соответствии с данным определением приспособленности целью этой искусственной эволюции будет как раз создание наилучших решений. Очевидно, эволюция – бесконечный процесс, в ходе которого приспособленность особей постепенно повышается. Принудительно остановив этот процесс через достаточно долгое время после его начала и выбрав наиболее приспособленную особь в текущем поколении, можно получить не абсолютно точный, но близкий к оптимальному ответ.

Такова вкратце идея генетического алгоритма (ГА). Перейдем теперь к точным определениям и опишем работу ГА более детально.

Для того чтобы говорить о генетическом наследовании, нужно снабдить имеющиеся существа хромосомами. В ГА хромосома – это некоторый числовой вектор, соответствующий подбираемому параметру, а набор хромосом данной особи определяет решение

задачи. Какие именно векторы следует рассматривать в конкретной задаче, решает сам пользователь. Каждая из позиций вектора хромосомы называется геном.

Определим понятия, соответствующие мутации и кроссинговеру в ГА.

Мутация – это преобразование хромосомы, случайно изменяющее одну или несколько ее позиций (генов). Наиболее распространенный вид мутаций – случайное изменение только одного из генов хромосомы.

Кроссинговер (в литературе по генетическим алгоритмам также употребляется название кроссовер или скрещивание) – это операция, при которой из двух хромосом порождается одна или несколько новых хромосом. В простейшем случае кроссинговер в ГА реализуется так же, как и в биологии. При этом хромосомы разрезаются в случайной точке и обмениваются частями между собой. Например, если хромосомы (1, 2, 3, 4, 5) и (0, 0, 0, 0, 0) разрезать между третьим и четвертым генами и обменять их части, то получатся потомки (1, 2, 3, 0, 0) и (0, 0, 0, 4, 5). Блок-схема ГА изображена на рис. 2. Вначале генерируется начальная популяция особей (индивидуумов), т. е. некоторый набор решений задачи. Как правило, это делается случайным образом. Далее необходимо смоделировать размножение внутри этой популяции. Для этого случайно отбираются несколько пар индивидуумов, производится скрещивание между хромосомами в каждой паре, а полученные новые хромосомы помещаются в популяцию нового поколения. В генетическом алгоритме сохраняется основной принцип естественного отбора – чем приспособленнее индивидуум (чем больше соответствующее ему значение целевой функции), тем с большей вероятностью он будет участвовать в скрещивании. Теперь моделируются мутации — в нескольких случайно выбранных особях нового поколения изменяются некоторые гены. Затем старая популяция частично или полностью уничтожается и переходят к рассмотрению следующего поколения. Популяция следующего поколения в большинстве реализаций генетических алгоритмов содержит столько же особей, сколько начальная, но в силу отбора приспособленность в ней в среднем выше. Теперь описанные процессы отбора, скрещивания и мутации повторяются уже для этой популяции и т. д.

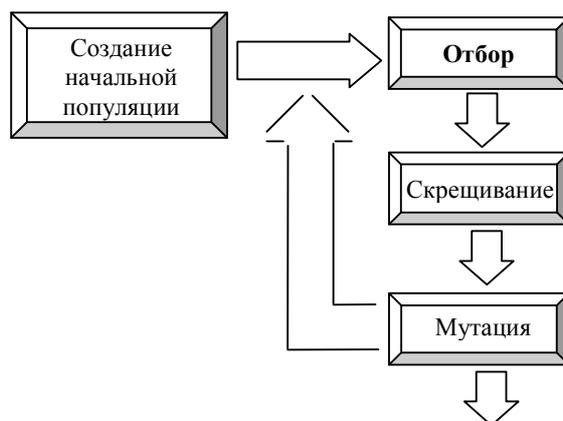


Рис. 2

В каждом следующем поколении возникают совершенно новые решения исходной задачи. Среди них будут как плохие, так и хорошие, но благодаря отбору число хороших решений будет возрастать. Эта процедура продолжается до тех пор, пока не найдется особь, гены которой представляют оптимальный набор параметров, при которых значение

целевой функции близко к максимуму или минимуму, либо равно ему. Останов работы ГА может произойти также в случае, если популяция вырождается, т. е. если практически нет разнообразия в генах особей популяции, либо если просто вышел лимит времени [2].

Может создаться впечатление, что ГА являются просто модификацией случайного поиска. Однако приспособленность была введена совсем не зря. Дело в том, что она непосредственно влияет на шанс особи принять участие в селекции. Выбирая каждый раз для скрещивания наиболее приспособленных особей, можно с определенной степенью уверенности утверждать, что потомки будут либо не намного хуже, чем родители, либо лучше их.

В соответствии с концептуальной моделью МУ к множеству оптимизируемых параметров относятся:

- 1) число процессоров обработки пакетов N ;
- 2) число секций общего поля памяти M ;
- 3) матрица распределения ресурсов по секциям памяти $B^P = [B_{mr}^P]_{m=1, \overline{M}; r=1, \overline{R}}$, где R –

число общих ресурсов.

Значение B_{mr}^P есть число элементов r -го ресурса, распределенных в m -ю СОП.

Теперь задачу определения структуры мультипроцессорной системы можно сформулировать следующим образом.

При заданных значениях Π : производительности процессоров обработки пакетов, объемов локальных памятей ПОП, стоимости модулей, составляющих структуру межсетевое устройства, определить такую оптимальную структуру МУ, чтобы выполнялись следующие условия по производительности $\Lambda(\Pi)$, средней задержке \bar{t}_3 и вероятности блокировки пакетов P_0^y :

$$\Lambda(\Pi) \geq \frac{1}{L} \sum_{y=1}^Y v_y \delta_y, \quad (1)$$

где L – средняя длина пакета в битах, v_y и δ_y – скорость, бит/с, и коэффициент использования (загрузки) y -го канала связи;

$$\bar{t}_3(\Pi) \leq t_3^d. \quad (2)$$

Здесь t_3^d – допустимое среднее время задержки пакета в МУ, с;

для каждого входящего канала $y = \overline{1, Y}$

$$P_0^y(\Pi) \leq P_0^d, \quad (3)$$

где P_0^d – допустимая вероятность блокировки и искажения пакета. Кроме того

$$\Phi(\Pi) \rightarrow \min_{\Pi}, \quad (4)$$

где Φ – аддитивная функция стоимости всех модулей, входящих в состав МУ.

Решение сформулированной задачи распадается на 2 этапа. На первом этапе на соответствующих моделях находятся зависимости $\Lambda(\Pi)$; $\bar{t}_3(\Pi)$; P_0^y .

На втором этапе решается оптимизационная задача, определяемая выражениями (1)–(4).

Исходными данными для работы процедуры оптимизации являются параметры генетического алгоритма:

Z – число особей в популяции,

χ_1, χ_2, χ_3 – порядок операции кроссовера для разных хромосом;

q – вероятность мутации гена,

$\epsilon, K_{\text{доп}}$ – критерии останова.

Каждая особь популяции описывается набором из трех хромосом:

1) N ;

2) $B_{11}^P, \dots, B_{MR}^P$;

3) M .

Алгоритм оптимизации включает следующие шаги:

Шаг 1. Выбрать начальную популяцию S_0 , включающую Z особей. Особь $s \in S_0$ включить в популяцию только при выполнении ограничений (1)–(3). Положить номер популяции $k: = 0$.

Шаг 2. Вычислить приспособленность каждой особи популяции $\Phi(s)$, $s = 1, \dots, Z$, и приспособленность популяции в целом

$$\Phi_k^* = \min \{ \Phi(s) \mid s = 1, 2, \dots, Z \},$$

где s – номер особи в популяции, который позволяет однозначно определить соответствующий этой особи генотип; $\Phi(s)$ – целевая функция задачи оптимизации.

Шаг 3. Селекция. Случайным образом выбрать родителей s_1, s_2 из популяции k для следующего потомства в соответствии с распределением вероятностей $P_1, \dots, P_s, \dots, P_Z$, где

$$P_s = \frac{\Phi(s)}{\sum_{s \in S_k} \Phi(s)}.$$

Шаг 4. Скрещивание. Построить хромосому потомка s'_i по s_1, s_2 при помощи оператора χ_i – точечного кроссовера применительно к i -й хромосоме.

Шаг 5. Мутация. Модифицировать s'_i , с вероятностью q заменив значение каждого гена потомка на противоположное.

Шаг 6. Формирование новой популяции, которая объединяет решения следующего поколения. Для этого Z раз выполнить операции алгоритма, начиная с шага 3, помещая каждую вновь полученную особь в следующую популяцию S_{k+1} . Затем положить номер новой популяции $k: = k + 1$ и перейти к шагу 7.

Шаг 7. Останов. Если приспособленность L последних популяций увеличилась меньше чем на ϵ , или если $k = K_{\text{доп}}$, то остановить процесс эволюции. В качестве решения задачи оптимизации взять наилучшую из найденных особей последней популяции.

СПИСОК ЛИТЕРАТУРЫ

1. Вишневикий В. М. Теоретические основы проектирования компьютерных сетей. М.: Техносфера, 2003.
2. Кутузов О. И., Сергеев В. Г., Татарникова Т. М. Коммутаторы в корпоративных сетях. Моделирование и расчет. СПб.: Судостроение, 2003.
3. Рутковская Д, Пилиньский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы/Пер. с польск. М.: Горячая линия – Телеком, 2004.

Within the framework of the decision of a problem of organization of an exchange given between network objects the concept of adjusted modular structure of the gateway device is developed. The offered modular technology, provides the directed formation of shape of the gateway device with the given set of properties and allows to expand its configuration. The basis of offered procedure is made by the specially developed genetic algorithm of optimization of structure of the gateway device on productivity.

The gateway device, multiprocessor system, parallel processing, optimization of structure of the device, genetic algorithm

УДК 621.399

A. E. Романов

МЕТОДИКА ЭЛЕКТРОННОГО МАКЕТИРОВАНИЯ ЭЛЕКТРИЧЕСКИХ ЖГУТОВ

Предложена методика электронного макетирования электрического жгута, выполняемая на электронной модели изделия и ориентированная на автоматизированное проектирование сложных технических изделий.

Электронное макетирование, САПР

В современных сложных технических системах на долю электрических жгутов приходится не более 2% от стоимости изделия, однако их влияние на функционирование систем, особенно на надежность, весьма существенно. По некоторым данным, не менее 50% отказов авиационной техники происходит из-за проблем с электрическими соединениями. Имеются сведения, что невыявленные причины многих аварий зачастую также связаны с электрическими соединениями (например, относительно недавно выявленный эффект Arc Tracking, или распространение дугового разряда в проводах с полиамидной изоляцией, приводящее к катастрофическому эффекту прогорания жгута). Современный электрический жгут является сложным и высокотехнологичным объектом. Особенно это относится к специальным жгутам, применяемым в авиации, кораблестроении, ракетной и космической технике, наземной боевой технике, современных транспортных средствах (высокоскоростные поезда, автомобили высокого класса и т.п.), нефтяной и газовой промышленности, в системах обмена данными. В настоящей статье излагается методика электронного макетирования электрического жгута, выполняемая на электронной модели изделия и ориентированная на автоматизированное проектирование сложных технических изделий.

Для облегчения процесса проектирования жгута, правильного выбора и заказа компонентов компанией Tусо Electronics была разработана программа HarnWare. Процесс проектирования в этой программе начинается с ввода топологии жгута, длин и электрической схемы. В процессе работы программа HarnWare позволяет выбрать правильные компоненты, определить маршруты проводящих линий, автоматически уточняет длины проводов, рассчитывает последовательность скрутки проводов и диаметры отдельных участков жгута (отдельно пучка проводов, проводов с экраном, проводов с экраном и защитной трубкой), автоматически генерирует полную спецификацию заказных компонентов, оценивает трудоёмкость монтажа и контроля. Использование программы HarnWare значительно снижает трудоёмкость проектирования

сложных электрических жгутов. Автоматический расчёт диаметров наборных жгутов с учётом толщины экрана и оболочки позволяет автоматически выбирать из базы данных термоусаживаемые компоненты подходящего размера и сразу вносить их в спецификацию.

Для корректной работы программы HarnWare конструктор должен задать положение жгута в пространстве, а топология жгута (разветвления, радиусы изгиба отдельных ветвей, точные длины плеч) может быть получена лишь в результате процесса макетирования. Процесс макетирования по сей день остается одним из наиболее трудоемких процессов при проектировании электрического жгута. Задача формулируется так: на изделии с ответными частями разъёмов прокладывается жгут по приемлемому пути. Оптимизируется длина жгута с учетом ограничений, которые накладываются зонами высокой температуры и электрическими параметрами, требованиями электромагнитной совместимости, допустимыми пространственными ограничениями.

В качестве альтернативы традиционному ручному макетированию предлагается электронное макетирование. Для электронного макетирования электрического жгута используется электронная трехмерная модель изделия, созданная в системе трехмерного моделирования, CATIA V.5. [1][2]. Основой новых технологий является электронный макет изделия. Он позволяет интегрировать все этапы проектирования сложного технического изделия, такие, как размещение оборудования, трубопроводов, электрических жгутов, например, на авиационном двигателе. Электронная модель используется всеми участниками проекта для выполнения параллельного проектирования [3]. В современных проектах эта модель не разрабатывается специально для макетирования, а является результатом предыдущих этапов разработки изделия. Для проектирования топологии жгута используется модуль Electrical Harness Installation, Electrical Library. В этом случае методика макетирования предусматривает следующие этапы:

1. Подготовить исходные данные – электронную трехмерную модель изделия, диаметры сечения проводов и электрическую схему будущего жгута, задать положения ответных частей разъёмов в 3D модели.

2. Определить по электрической схеме предварительную топологию жгута, разветвления, положения ключей на разъемах.

3. Создать электронные модели разъёмов будущего жгута, с использованием модуля Electrical Librarian и Part Design системы CATIA. Подробное и точное изображение не требуется, достаточно соблюсти реальные размеры разъёмов.

4. Построить модель жгута, используя модуль Electrical Harness Installation, в соответствии с предварительной топологией, полученной на этапе 2. Построенную электронную модель жгута проложить на электронной модели изделия. Жгут прокладывается по оптимальной трассе. Когда жгут проложен, соединён всеми разъёмами с ответными частями на изделии, конструктор получает информацию о длинах его плеч. На этом этапе вычислить диаметры плеч жгута в программе HarnWare, которая на основе информации о длинах плеч (полученной из модели в CATIA) и сечении проводов, вычисляет диаметры всех участков жгута.

5. Добавить информацию о диаметрах плеч в эскиз 3D-модели CATIA и ещё раз проверить, проходит ли жгут по той трассе, по которой его проложили. В случае, когда из-за диаметра какого-либо участка жгут не проходит, например, через какое-то отверстие или трубку, можно повторить весь процесс и попытаться проложить его по новой трассе или предложить конструктору изделия внести определенные изменения в конструкцию.

В заключение хочется заметить, что практическая реализация методики электронного макетирования электрических жгутов, разрабатываемое и внедряемое компанией Bue Pitron, существенно сокращает трудоёмкость макетирования и расход дорогих материалов на макеты жгутов.

СПИСОК ЛИТЕРАТУРЫ

1. Голдовский П. С., Лесников Е. И. САТИА V5 покоряет три стихии. “САПР и графика”, 2002 г., №6
2. Бубнов А. А. САПР в судостроении. – статья в журнале “САПР и графика”, 2000 г., №5
3. Захаров В. Программное обеспечение САТИА V5 для проектирования двигателей. – статья в журнале “САПР и графика”, 2002 г., №9

A. E. Romanov

METHODS OF THE ELECTRONIC PROTOTYPING ELECTRIC HARNESSSES

Offered methods of the electronic prototyping electric harness, executed on electronic model of the product and oriented on using for automatic designing complex technical products.

Electronic prototyping, CAD

УДК 37.8 К40

Н. В. Комолова

О МЕТОДОЛОГИИ РАЗРАБОТКИ СТРАТЕГИИ УПРАВЛЕНИЯ ПРЕДПРИЯТИЯМИ ОТРАСЛИ ПЕЧАТИ

В предложенной методологии разработки стратегии управления предприятиями отрасли печати уточнены понятия потенциала предприятия и показатели эффективности работы. Предложена модель потребностей общества в печатной продукции. Определена роль автоматизированных информационно-управляющих систем в едином информационном пространстве предприятия.

Управление, информация, автоматизированные информационно-управляющие системы, модель потребностей

Сегодняшний бизнес – это мгновенная реакция на любые изменения, быстрая адаптация к постоянно меняющимся условиям и конъюнктуре рынка. Поэтому решения по управлению средними и крупными предприятиями и корпорациями должны быть точными, мощными, гибкими, масштабируемыми и оптимальными. Производство печатных средств информации является многогранным, сложным процессом. Такими же многогранными являются и требования рынка, предъявляемые к печатной продукции, поэтому возникает потребность в выработке стратегии управления предприятиями отрасли печати, а в частности, методологии.

Предлагается следующая методология выработки стратегии управления предприятиями отрасли печати:

1. Постановка задачи. Задание потенциала полиграфического предприятия.

Уточнение стоимости или гудвилла фирмы (предприятия, корпорации), профиля деятельности, ассортимента выпускаемой продукции, места расположения офиса и производственных помещений, занимаемой предприятием площади, численности персонала предприятия, оснащение фирмы современным полиграфическим и другим оборудованием, возможность массового обслуживания потока заказчиков, возможность вовлечения в инвестиционные проекты; поддержание высокой рентабельности предприятия; возможность создания де-

шевой и качественной продукции, стоимость основных средств, стоимость основных производственных фондов, стоимость основных непроизводственных фондов, стоимость оборотных средств, стоимость оборотных производственных фондов, стоимость фондов обращения, стоимость материальных ресурсов, стоимость денежных средств, стоимость финансовых ресурсов, рыночная стоимость предприятия, прибыль, уровень фондоотдачи, капиталотдача, ресурсоемкость, фондоемкость, эффективность использования активов, материалоемкость, капиталоемкость, трудоемкость, энергоемкость, топливоемкость, рентабельность, ликвидность, платежеспособность, производительность труда, объем производства продукции (работ, услуг); валовой доход, чистая продукция, уровень автоматизации, общий тираж книг, газет и журналов; количество различных видов печатных машин; количество различных видов печатной продукции; объем изготовленной продукции. Отдельным вопросом оценки состояния фирмы следует считать автоматизированные информационно-управляющие системы, внедренные или подлежащие внедрению на предприятиях отрасли печати.

2. Построение математической модели.

Для изучения реального производственного процесса предприятий отрасли печати необходимо проведение математического моделирования процессов полиграфического производства, производственных потоков, процессов массового обслуживания потока заказчиков, моделирование рекомендаций по управлению производством, вырабатываемых автоматизированными информационно-управляющими системами (АИУС), внедренными или подлежащими внедрению на предприятии отрасли печати.

Отрасль печати развивается во времени. Пусть

x_i – неизвестные стратегии управления, придающие устойчивость отрасли и обеспечивающие максимальные прибыли;

ξ_i – некоторые случайные процессы с заданными траекториями, описываемые флуктуациями внешней среды. Тогда стратегии управления являются решениями уравнения

$$\frac{dx_i}{dt} = F_i(x_1, \dots, x_n, \xi_1, \dots, \xi_m) \quad (1)$$

Для упрощения уравнения (1), можно линеаризовать ξ_i , чтобы получить модель, определенную стохастическими дифференциальными уравнениями:

$$dx(t) = f(x)dt + \sum_{j=1}^l g_j(x)dw_j \quad (2)$$

где функции $f(x)$ определяют неслучайный вклад в динамику; $g_j(x)$ – функции, w_j определяют поток флуктуаций (стандартный процесс Wiener).

Описанная таким образом общая модель стратегий управлений является классической физической модель, слишком абстрактной и зависящей от внешних данных, поэтому конкретные стратегии управления необходимо разрабатывать с помощью имитационной модели.

3. Выбор показателей и критериев эффективности решения задачи.

В качестве показателей эффективности работы предприятий отрасли печати могут быть выбраны различные показатели, например, повышение капитализации (стоимости

имущественного комплекса предприятия или гудвилла) фирмы, повышение стоимости основных средств, прибыль, уровень фондоотдачи, капиталоемкость, ресурсоемкость, фондоемкость, эффективность использования активов, материалоемкость, капиталоемкость, трудоемкость, рентабельность, ликвидность, платежеспособность, производительность труда, объем производства продукции (работ, услуг).

4. Статистическое моделирование на имитационной модели.

Статистическое моделирование представляет собой метод получения с помощью ЭВМ статистических данных о процессах, происходящих в моделируемой системе [1]. Статистические данные обрабатываются и классифицируются с использованием методов математической статистики. Проведение имитационных экспериментов с моделями сложных систем позволяет пользователям-разработчикам моделей оценить различные стратегии по управлению сложными системами.

5. Оценка эффективности результатов.

Результаты моделирования на статистической модели показали, что только автоматизированные системы управления должны обеспечить высокую эффективность выполнения самых разнообразных процессов и предложить решения по управлению производством в различных отраслях народнохозяйственного комплекса, в том числе и в такой важной отрасли, от которой во многом зависит экономический рост общественного производства, как отрасль печати.

6. Рекомендации по внедрению или модернизации имеющейся АИУС. Внедрение результатов работы.

На основе модели выработки стратегии управления предприятиями отрасли печати определяются структура, алгоритмы и параметры АИУС, выбираются аппаратно-программные средства реализации системы.

В настоящее время для поддержки совместной работы людей уже недостаточно привычных средств общения. Все чаще решение бизнес-задач требует применения коммуникационных технологий, работающих в режиме реального времени. Необходимо также интегрировать в программное обеспечение поддержку групповой работы широкого спектра способов общения: электронную почту; телефон; мгновенные сообщения, instant messaging, IM; SMS; видео- и Web-конференции.

Предприятия отрасли печати выпускают продукцию, пользующуюся спросом, поэтому необходимо обосновать методы выбора стратегии управления предприятиями отрасли печати. Так, например, емкость рынка печатных машин составляет 8,2 млрд. €. Мировой коммуникационный рынок в 2003 г. составил 1820 млрд € (из них телекоммуникации составляют 1050 млрд €, печать – 500 млрд €, телевидение – 185 млрд €, Интернет – 50 млрд €, радио – 35 млрд €).

Традиционно допечатные процессы делятся на три области: набор, т.е. подготовка текста, его форматирование; репродуцирование изобразительных и графических оригиналов, изготовление цветоделенных изображений для многокрасочной печати; монтаж и изготовление печатных форм, т.е. объединение текстовых, иллюстративных и графических составных частей в одно целое на полосах, на печатных листах, а также изготовление печатной формы как носителя информации для печатной машины.

Печатью называется процесс переноса краски на бумагу (или на другой запечатываемый материал) посредством печатной формы. Можно выделить 4 способа печати, которые в зависимости от используемой печатной формы подразделяются:

– на высокую печать, в которой верхние участки печатающих элементов (буквы, линии, растровые точки и т. д.) выступают над поверхностью формы и расположены в одной плоскости.

К высокой печати относится и флексографская печать, которая в основном используется для печати на упаковке. В типографской книжной печати используется жесткая печатная форма (из металлического сплава), во флексографской печати – мягкая, эластичная печатная форма;

– глубокую печать; в которой печатающие элементы на формном материале, наоборот, углублены.

– плоскую печать, в которой печатающие и пробельные элементы располагаются в одной плоскости. Главной разновидностью плоской печати является офсетная печать, являющаяся в настоящее время доминирующим способом печати.

Послепечатная обработка продукции охватывает такие операции, как резка, фальцовка, подборка листов и переплетные работы, являющиеся важными составными элементами конечного этапа производства печатной продукции.

В полиграфическом производственном процессе печатные машины занимают центральное место. Печатные машины изготавливаются как для листовой, так и для рулонной печати.

Положительно влияют на полиграфическую промышленность компьютерные технологии. Уже широко используются системы «Компьютер – фотоформа» (Computer to Film), «Компьютер – печатная форма» (Computer to Plate) и «Компьютер – печатная машина» (Computer to Press), и их распространение продолжается.

При интеграции допечатных процессов, печати и послепечатной обработки центральное значение имеют стандартизированные форматы данных, так как они обеспечивают реализацию «рабочего потока» Workflow.

Свыше 40 известнейших производителей оборудования и систем для полиграфии работают в рамках международной кооперации над тем, чтобы с помощью так называемой концепции СІРЗ (СІРЗ – это международная кооперация в областях допечатных, печатных и послепечатных процессов) разработать стандарт для цифровых процессов выпуска печатной продукции.

Предлагается следующая модель потребностей общества в печатной продукции.

Потенциальные покупатели продукции отрасли печати, обладающие капиталом, достаточным для совершения покупки, распределены на некоторой территории земного шара, обитая в городах, населенных пунктах и т.д.

Пусть $\rho(\omega)$ – плотность распределения потенциальных покупателей, тогда:

$$\int_{\omega_1}^{\omega_2} \rho(\omega) d\omega - \text{количество потенциальных покупателей на некоторой территории};$$

$\rho(p)$ – плотность распределения платежеспособности покупок, тогда

$$\int_{p_1}^{p_2} \rho(p) dp - \text{прибыль от продажи полиграфической продукции за интервал времени,}$$

на интервале площади и $P = \int_{p_1}^{p_2} \int_{\omega_1}^{\omega_2} \int_{t_1}^{t_2} \rho(p, \omega, t) dp d\omega dt$ – прибыль от продажи полиграфической продукции.

Экстремальная задача формализуется следующим образом:

$$P \rightarrow \inf(\sup)$$

Каждое предприятие, конечно же, в своей работе стремится к максимизации прибыли. Поступление денег и прибыль образует

$G - CW - complex$ клеточное пространство; при этом:

P_0 – одноклеточное пространство, которое образуют единичные потребители и производители;

P_1 – двуклеточное пространство, которое образуют группы потребителей и производителей;

P_n – $(n - 1)$ клеточное пространство корпораций, холдингов, консорциумов и т.д.

Тогда, $P_n \rightarrow P_{n-1} \rightarrow \dots \rightarrow P_0$.

Пусть

X – базисное пространство прибылей, возможных при совершении коммерческих сделок;

X / G_0 – фактор-группа полученной прибыли в базисном пространстве;

\tilde{X} – накрывающее пространство;

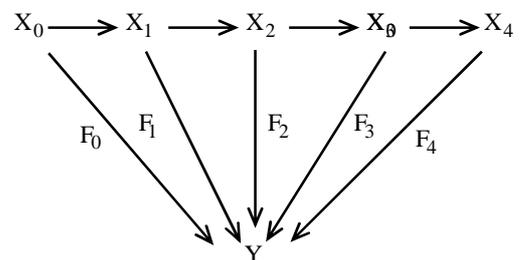
\tilde{X} / \tilde{G} – фактор-группа полученной прибыли в накрывающем пространстве;

Тогда коммутативная диаграмма получения прибыли при предложении товаров потребителям одной фирмой выглядит следующим образом:

$$\begin{array}{ccc} \tilde{X} & \rightarrow & \tilde{X} / \tilde{G}_0 \\ \downarrow & & \downarrow \\ X & \rightarrow & X / G_0 \end{array}$$

Коммутативная диаграмма процесса образования капитала Y от процесса предложений полиграфической продукции и последующих предложений имеет вид [3, 4]:

Информация является источником для принятия разнообразных решений [2]. В процессе управления предприятием его руководитель и его заместители сталкиваются с проблемой наличия огромных информационных потоков, влияющих на принятие управленческих решений. АИУС являются элементом единого информационного пространства предприятия. В АИУС должны быть реализованы функции интеллектуального управления коммуникациями с возможностями сотрудников в конкретной ситуации.



Для принятия эффективного решения по управлению предприятием на помощь руководителям приходят АИУС, реализованные на предприятии. Однако, они только рекомендуют конкретные решения, принятие решения – прерогатива руководителя. На рисунке показана структура отрасли печати, имеющей единое информационное пространство предприятий.



Выводы: предложена методология выработки стратегии управления предприятиями отрасли печати. Установлено, что сегодня принятие управленческого решения превращается в верхушку айсберга, в основании которого лежит решение множества задач, связанных со сбором и организацией информации. В связи с этим принятие эффективных управленческих решений на управление невозможно без внедрения АИУС. Разработана структура отрасли печати, имеющей единое информационное пространство предприятий.

СПИСОК ЛИТЕРАТУРЫ

1. Советов Б. Я., Яковлев С. А. Моделирование систем: Учеб. для вузов. -2-е изд., перераб. и доп., М.: Высш. шк., 1998.
2. Советов Б. Я., Цехановский В.В. Информационные технологии: Учеб. Для вузов. – М.: Высш. шк., 2003. – 263 с.: ил.
3. Luck, Wolfgang. Transformation groups and algebraic K-theory, 1989, Lecture Notes in Mathematics vol. 1408, pages 444.
4. Luck, Wolfgang. L2-Invariants: Theory and Applications to Geometry and K-Theory. Publisher: Springer Verlag, 2002

N. V. Komolova

ABOUT THE METHODOLOGY OF STRATEGY OF DEVELOPMENT PRINT PUBLISHING PRODUCTION

An approach to building the methodology of strategy of development the print publishing. The model of necessity of print publishing production are offered. The role of the automatic information-management system in quality information space on the factory have described.

Management, information, automatic information-management system, model of wants

СЕТЕВОЕ УПРАВЛЕНИЕ АППАРАТУРОЙ ВЧ-СВЯЗИ

Рассматриваются варианты сетевого обмена данными между приложениями, разработанными с использованием языка программирования высокого уровня Java.

Java, сокет, RMI, компьютерная сеть

Аппаратура высокочастотной связи (ВЧ-связи) позволяет организовать диспетчерские телефонные каналы для оперативного руководства в местах, где обычные виды связи невозможно использовать либо стоимость эксплуатации слишком велика. Помимо этого она позволяет производить удаленный контроль разнообразной аппаратуры, объектов электроэнергетики.

Современные сетевые технологии позволяют передавать и принимать файлы или данные, выполнять сложнейшие математические расчеты или просто напрямую общаться с собеседником, находящимся на другой стороне земного шара. При этом часть программ строится таким образом, что на рабочую станцию устанавливается программа, в функции которой входит только прием запросов и данных от пользователя; передача введенных данных на обработку; прием и вывод результатов обработки пользователю. Все сложные процедуры обработки выполняются на сервере, что позволяет на рабочем месте использовать аппаратное обеспечение, не приспособленное к решению сложных задач. В качестве примера можно привести кассовые терминалы в супермаркетах: кассир подносит товар со штрих-кодом к сканеру, терминал, получая номер от сканера, передает его на центральный сервер, где происходит поиск наименования товара, его цены, и, при необходимости, дополнительной информации. Результат запроса передается обратно на терминал кассира, выводится на дисплей и чек.

Используя JAVA в качестве основного языка программирования, разработчики могут не задумываться, на каком аппаратно-программном комплексе будут запускаться программы, просто там должна стоять JAVA-машина (Java Runtime Environment – JRE) и, при необходимости, обеспечен интерфейс взаимодействия с внешними устройствами.

Как правило, при использовании JAVA сетевой обмен начинается с создания сокета на ближней стороне и открытия сокета на удаленной стороне.

В данной статье ставится задача определения наиболее удобного способа для реализации обмена между клиентом и сервером через сеть как для программ, созданных одной командой программистов, так и для сторонних разработчиков. Для простоты сравнения рассмотрим только способы реализации сетевого обмена.

Будут рассмотрены следующие варианты обмена: байтовый, заранее определенными объектами пользователя и используя механизм RMI. Основные параметры, по которым будет происходить сравнение:

- простота создания и удобство использования;
- относительные требования к ресурсам;
- возможность обмениваться данными с программами других разработчиков.

Создание программ с использованием байтового обмена является самым простой и быстрый из всех возможных способов передачи данных по сети. Байтовый обмен – это простой ввод-

вывод: байты передаются из одной системы в другую, байт за байтом. При этом разработчик программы обязан сам отслеживать, чем является каждый следующий байт принятой информации. Это может быть число, буква (в однобайтовой кодировке) или часть более сложной структуры. Так, например, для передачи одной строки необходимо получить ее длину, привести строку к массиву из байт, сформировать признак типа данных (строка) и только после этого можно начинать передачу. При этом нет никакого механизма для обеспечения гарантированной доставки байта от одной системы к другой. Конечно, можно предусмотреть подтверждение доставки, но это усложнит всю процедуру и снизит скорость передачи данных.

Использование объектов, определяемых пользователем, позволяет не задумываться о том, как будет происходить передача данных и что именно означает каждый байт. Если для передачи используются не простые типы данных, известные *каждой* JRE, то необходимо создать специальное описание объекта, предоставляемое и клиентской, и серверной сторонам. Для изменения параметров объекта, необходима повторная компиляция как серверного, так и клиентского приложений. В противном случае при попытке осуществить обмен возникнет ошибка. При этом отсутствует необходимость в каких либо дополнительных механизмах установления соединения между клиентом и сервером. По сути, это расширение байтового обмена.

Использование RMI (Remote Method Invocation) позволяет разработчикам строить сложные системы, взаимодействующие через сеть [см. лит.]. Данный механизм был специально разработан для создания распределенных клиент-серверных приложений. RMI позволяет из одной работающей JAVA-машины вызвать функцию для компонента RMI, расположенного на другой JAVA-машине. Для этого необходимо определить свойства компонента и выполнить реализовать его. С помощью *языка описания интерфейса* (IDL) описать интерфейс объекта, с которым будет происходить взаимодействие, затем создать компонент (используя один из языков программирования), поместить его в нужную точку сети (на один или несколько серверов), зарегистрировать объект у менеджера объектов, после чего он готов к использованию. Данный механизм позволяет использовать удаленные сетевые ресурсы напрямую.

Требования к ресурсам при использовании байтового обмена минимальны как при формировании запроса, так и при приеме. При добавлении обработки принятых байт, в зависимости от логики программы, требования могут повышаться, но в целом нагрузка на систему заметно не возрастает.

При использовании объектов пользователя всегда выделяется память, вне зависимости от того, необходимо ли данное поле или нет, и при пересылке по сети объект передается целиком. Достоинство такого метода – весьма эффективный способ обмена, хотя при этом в целом возрастает нагрузка на систему и сеть.

RMI предъявляет довольно серьезные требования к системе, предоставляющей RMI-компоненты для совместного использования. Чаще всего система поддерживает помимо самого объекта еще и специальную базу данных, хранящую его описание и доступные для этого объекта методы. При грамотной реализации методов объекта сетевой обмен не намного превышает минимально необходимый для передачи всех данных.

Возможность обмена данными с программами других разработчиков при байтовом обмене – очень сложная для реализации задача. Малейшее изменение программы, связанное с сетевым обменом (например, порядок получения или разбора байт), может привести к полной потере работоспособности.

Использование в качестве единицы обмена определенного пользователем объекта позволяет изменять последовательность получения данных при использовании объектов одного типа. Для взаимодействия сторонним разработчикам необходимо получить откомпилированный класс с описанием объекта и включить его в свой проект. Если в обмене участвуют объекты различных типов, достаточно сложно оценить, что именно сейчас пришло. В таком случае рекомендуется все привести к одному типу.

При использовании RMI подобные сложности отсутствуют. Любой разработчик, получив от брокер-менеджера описание объекта, может подключиться к этому объекту и вызывать его методы. При этом, как правило, входные и выходные данные являются простыми типами. Разработчик серверной части может легко менять всю логику поведения метода объекта при условии неизменности входных и выходных данных.

Рассмотренные выше методы сетевого обмена при создании программного обеспечения только с использованием JAVA позволяет сделать вывод, что наиболее универсального способа сетевого обмена не существует.

В более простых приложениях, разрабатываемых одной командой программистов, если не планируется возможность обмена со сторонними программами, для упрощения разработки и снижения нагрузки на систему можно порекомендовать использовать простейшие объекты, определяемые разработчиками. Для очень сложных сетевых приложений, требующих хорошей интеграции с существующими и разрабатываемыми приложениями, наилучшим вариантом будет использование RMI. Байтовый обмен может представлять интерес как основа создания собственных классов для передачи данных по сети.

Оценка всех рассмотренных ранее методов сетевого обмена была выполнена в фирме ООО «НПФ Модем». Программа разбита на серверную и клиентскую части. При этом серверная часть осуществляет прием управляющих команд, разбор команды для анализа, кому предназначена команда, и пересылку, при необходимости, командной последовательности в аппаратуру. Исходя из требований простоты разработки, последующей работы необученного персонала и снижения нагрузки на систему в целом был реализован обмен с использованием пользовательских объектов.

Командная последовательность полностью формируется на клиентской стороне и перед отправкой серверу помечается соответствующим образом. Серверная часть принимает ответ от аппаратуры и передает его клиентской части.

Данный подход позволяет управлять оборудованием, находящимся на значительном расстоянии, не вдаваясь в подробности, каким именно образом осуществляется связь, главное требование – между клиентом и сервером должно быть соединение по протоколу TCP/IP.

СПИСОК ЛИТЕРАТУРЫ

1. Elliott Rusty Harold, Java network programming, O'Reilly, Sebastopol, 2004 г.

D. Haysov

HIGH VOLTAGE HARDWARE NETWORK CONTROL

In article analyses variants for exchange over network data between software, making with java high level program language.

Java, socket, RMI, computer network

МЕТОДЫ ОПРЕДЕЛЕНИЯ КОНСИСТЕНТНОГО СОСТОЯНИЯ В СИСТЕМАХ С САМОВОССТАНОВЛЕНИЕМ

Описываются основные методы определения консистентного состояния распределенной системы, основанные на алгоритмах определения процессами контрольных точек. Предложен коммуникационно-вынужденный подход при взятии контрольных точек и представлены процедуры для отправки и приема сообщений в данном механизме.

Консистентное состояние системы, локальные и глобальные контрольные точки, сообщения-сироты, недоставленные сообщения, коммуникационно-вынужденный подход

Модели восстановления распределенных систем основаны на процедурах мониторинга, измерений параметров и классификации с учетом предыстории работы отдельных узлов или подсистем, реконфигурации и при необходимости направленной функциональной деградации. Важнейшим механизмом, поддерживающим реконфигурацию, является механизм определения контрольных точек и откатов. В течение времени выполнения приложения алгоритм контрольных точек вычисляет и сохраняет глобальные контрольные точки (глобальные состояния) на устойчивом носителе. В случае возникновения ошибки алгоритм откатов восстанавливает вычисление с наиболее позднего глобального состояния, с которого это вычисление может быть возобновлено.

Глобальная контрольная точка (ГКТ) состоит из двух частей: (1) множества локальных КТ (ЛКТ), каждая из которых определена для каждого процесса, и (2) набора сообщений (состояний канала передачи) для каждого канала связи. Неформально, ГКТ является консистентной, если вычисление могло пройти через нее.

Известны многие алгоритмы определения контрольных точек. Их можно разделить на 2 класса согласно способу определения ЛКТ: координированные и некоординированные.

Координированные алгоритмы. Класс координированных алгоритмов подразумевает синхронное определение ЛКТ и вычисление соответствующих состояний канала таким образом, что результирующие ГКТ являются всегда консистентными. Впервые такого рода координированный алгоритм был предложен Чанди и Лампортом [1] в 1981 г. Они рассматривали распределенную систему в виде сильно-связанного, конечно-ориентированного графа, в котором каждая вершина является машиной, представляющей один из процессов в системе, и каждая дуга является однонаправленным FIFO – буфером, которому соответствует канал передачи данных между процессами.

Распределенное вычисление состоит из последовательности так называемых элементарных событий. Каждое элементарное событие происходит на одной из машин: оно изменяет ее состояние, принимая одно сообщение во входной буфер или посылая одно сообщение в выходной буфер этой машины. В описании алгоритма, который был назван алгоритмом снимка (снимок есть ГКТ), каждая машина, каждое элементарное событие и каждое сообщение являлись либо белым, либо красным. Каждое элементарное событие получает цвет машины, которая производит это событие, каждое сообщение также получает цвет машины, которая отправила это сообщение. Определение снимка в их работе описывается как переход системы из со-

стояния S_0 в состояние S_1 . В состоянии S_0 все машины и все сообщения белые. В процессе реализации алгоритма каждая машина превращается из белой в красную только однажды, т. е. локальная хронология каждой машины есть последовательность белых событий и последующая последовательность красных. В конечном счете, все машины и все события будут красными. Состояние S_1 – состояние, в котором все машины и все сообщения красные.

Красная машина должна делать запись для каждого входного буфера строки принятых сообщений включая последнее белое. Поэтому авторы расширили набор сообщений еще одним специальным, называемым «маркером» («красным письмом»), с соглашением, что в каждом буфере сообщения (если они есть), которые предшествуют маркеру, считаются белыми, а сообщения (если они есть), следующие за маркером – считаются красными; подобно другим сообщениям, маркеры участвуют в FIFO-режиме буферов, и после превращения в красную каждая машина посылает по каждому из выходных буферов маркер перед посылкой дальнейших сообщений.

При приеме красного сообщения белая машина посылает одно красное письмо по каждому из своих выходных буферов и становится красной. Если принимающая машина уже была красной, она остается таковой и не посылает дополнительных красных писем.

Так как каждому красному сообщению предшествует маркер, машины становятся красными после принятия первого маркера, являясь еще белыми. Алгоритм снимка инициализируется одной машиной, становящейся красной (и, соответственно, при посылке маркера по каждому из ее выходных буферов). Так как каждая машина становится красной только однажды, каждый буфер несет точно один маркер. Следовательно, каждой машине известно, когда ее локальная информация снимка будет собрана, т. е. когда по каждому входному буферу был принят маркер. Локальная информация снимка может тогда быть послана центральному пункту (вероятно, только инициатору), где она будет сохранена и в дальнейшем проанализирована [1].

В дальнейшем данный алгоритм был неоднократно переосмыслен и дополнен, но основная идея, заложенная в нем, – синхронизация при определении ГКТ – осталась.

Главный недостаток такого подхода – дополнительная синхронизация, требующая наличия управляющих сообщений, которые могут замедлять или даже замораживать вычисление. Главное преимущество состоит в том, что последняя вычисленная ГКТ (так как она является консистентной) является единственной, которая должна сохраняться на устойчивом носителе. Кроме того, при отказе, реализация алгоритма восстановления состоит в считывании ГКТ и возобновлении процессов (вычисления) от нее [3],[4].

Некоординированные алгоритмы. В классе некоординированных алгоритмов процессы сохраняют ЛКТ на устойчивом носителе независимо друг от друга. Кроме того, они также регистрируют сообщения на устойчивом носителе так, что состояние каналов связи также протоколируется. В случае отказа алгоритм отката вычисляет консистентное глобальное состояние из ЛКТ и записанных сообщений. Способ регистрации сообщений называется ориентированным на отправителя (ориентированным на получателя), если сообщения регистрируются их отправителями (получателями). Возможны 2 метода регистрации сообщений. В случае пессимистической регистрации сообщения регистрируются на устойчивом носителе во время их посылки (получения). Это может нанести существенный вред при

безотказных вычислениях, так как каждая отправка (получение) влечет за собой дополнительный ввод-вывод. В связи с этим были предложены оптимистические методы регистрации. В этом случае посланное (полученное) сообщение сохраняется во временной памяти, и впоследствии переносится на постоянный носитель, когда процесс берет ЛКТ (поскольку все сообщения сохранены, устойчивый носитель содержит все недоставленные сообщения) [3].

Коммуникационно-вынужденный подход к определению консистентного состояния [2]. При таком подходе управляющая информация переносится посредством сообщений самих приложений. Когда рассматривается глобальная контрольная точка G , две категории сообщений являются особо важными: сообщения, которые были доставлены, но не были отправлены в G – сообщения-сироты, и сообщения, которые были посланы, но не были получены в G – недоставленные сообщения.

Глобальная контрольная точка является консистентной, если не существует сообщений-сирот по отношению к ней. Консистентность глобальной контрольной точки является важным фактором, так как она определяет глобальное состояние, через которое прошло распределенное вычисление или через которое оно могло пройти. Рассмотрение только консистентных глобальных контрольных точек в некоторых случаях является недостаточным для определенного рода приложений. В таких приложениях недоставленные сообщения по отношению к некоторой глобальной контрольной точке должны быть записаны, так как их игнорирование может привести к ситуации, где они будут потеряны. Когда явно рассматриваются недоставленные сообщения в ГКТ, получаем состояние системы (вычислений). Такое состояние состоит из ГКТ G и множества M_G сообщений, являющихся недоставленными в ГКТ G (записанные сообщения). По отношению к некоторому состоянию системы недоставленные и незаписанные сообщения (не входящие в M_G) называются пропущенными. Состояние является консистентным, если ассоциированная с ним ГКТ консистентна (нет сообщений-сирот) и все недоставленные сообщения зарегистрированы в этом состоянии (нет пропущенных сообщений).

Локальная контрольная точка – это записанное состояние процесса. Когда процесс P_i определяет локальное состояние как локальную контрольную точку, говорят, что P_i берет ЛКТ (на уровне операционной системы это означает, что P_i сохраняет свой контекст на устойчивом носителе).

Сообщение m , посланное процессом P_i процессу P_j , называется сообщением-сиротой по отношению к упорядоченной паре ЛКТ $(c_{i,x}, c_{j,y})$, если $(\text{del}(m) \in c_{j,y}) \wedge (\text{send}(m) \notin c_{i,x})$, т. е. если событие доставки сообщения m принадлежит $c_{j,y}$, в то время как соответствующее событие отправки не принадлежит $c_{i,x}$.

Сообщение m , посланное процессом P_i процессу P_j называется недоставленным сообщением по отношению к упорядоченной паре ЛКТ $(c_{i,x}, c_{j,y})$, если $(\text{send}(m) \in c_{i,x}) \wedge (\text{del}(m) \notin c_{j,y})$, т. е. если событие отправки сообщения m принадлежит $c_{i,x}$, в то время как соответствующее событие доставки не принадлежит $c_{j,y}$ [2].

Сообщение называется записанным, если его содержимое было записано на устойчивом носителе информации. Сообщение является пропущенным по отношению к упорядоченной паре ЛКТ $(c_{i,x}, c_{j,y})$, если оно не записано и является недоставленным по отношению к данной паре ЛКТ. Глобальная контрольная точка (ГКТ) – это набор ЛКТ, по одной на каждый процесс. ГКТ

является консистентной, если по отношению к любой паре ее ЛКТ нет сообщений-сирот. Состояние S – это пара, состоящая из ГКТ G и множества M_G записанных сообщений, которые являются недоставленными по отношению к ЛКТ из G . Множество сообщений M_G является консистентным, если оно содержит все недоставленные сообщения по отношению к ЛКТ из G .

Определение 1. Состояние $S = (G, M_G)$ является консистентным, если:

- ГКТ G является консистентной (нет сообщений-сирот);
- множество M_G является консистентным (нет пропущенных сообщений).

Описание поставленной задачи. Для заданного распределенного вычисления и набора основных контрольных точек проблема состоит в проверке выполнения двух свойств:

(P1) $\forall c_{i,x} : \exists G = (\dots, c_{i,x}, \dots)$ такая, что состояние (G, M_G) является консистентным, т. е.:

(P11) G является консистентной (нет сообщений-сирот);

(P12) M_G является консистентным (нет пропущенных сообщений);

(P2) $\forall c_{i,x} : \text{возможно ассоциировать «на лету» с } c_{i,x} \text{ консистентное состояние, к которому принадлежала бы КТ } c_{i,x}$.

Методом выполнения условия (P11) является необходимость взятия вынужденных КТ перед доставкой сообщений, которые могут вызвать нарушение выполнения данного условия. Другой способ состоит в задержке при доставке сообщения, которое может вызвать нарушение условия (P11). Но использование данного подхода может вызвать нарушение в доставке сообщений, если процессы не берут достаточное число основных контрольных точек.

Способом выполнения условия (P12) является запись достаточного числа сообщений, а именно тех, чья незапись может привести к нарушению данного условия. Другой подход подразумевает задержку при взятии основных контрольных точек, но в некоторых ситуациях такой подход может предотвратить «взятие» некоторых основных ЛКТ.

Способом выполнения условия (P2) является ассоциирование с каждой консистентной ГКТ $G = (\dots, c_{i,x}, \dots)$ временной метки, которая известна процессу P_i , когда он берет ЛКТ $c_{i,x}$.

Временные метки. Опишем коммуникационно-вынужденный протокол состояний, основанный на механизме временных меток. С каждой ЛКТ $c_{i,x}$ ассоциируется временная метка Лампорта $c_{i,x}.t$, и каждое сообщение m содержит временную метку $m.t$.

Рассмотрим следующие определения для интервала контрольной точки $c_{i,x}$:

Множества:

$\text{sent}_{i,x} = \{m \mid \text{send}(m) \text{ произошло после } c_{i,x-1} \text{ и перед } c_{i,x}\}$

$\text{rec}_{i,x} = \{m \mid \text{del}(m) \text{ произошло после } c_{i,x-1} \text{ и перед } c_{i,x}\}$

$\text{sent_nr}_{i,x} = \{m \mid m \in \text{sent}_{i,x} \text{ и } m \text{ – незаписано}\}$

$\text{rec_nr}_{i,x} = \{m \mid m \in \text{rec}_{i,x} \text{ и } m \text{ – незаписано}\}$

Значения:

$\text{Min_sent}_{i,x}$ = наименьшая временная метка в сообщениях в $\text{sent}_{i,x}$, если $\text{sent}_{i,x} \neq \emptyset$, иначе $+\infty$.

$\text{Max_rec}_{i,x}$ = наибольшая временная метка в сообщениях в $\text{rec}_{i,x}$, если $\text{rec}_{i,x} \neq \emptyset$, иначе $-\infty$.

$\text{Max_sent_nr}_{i,x}$ = наибольшая временная метка в сообщениях в $\text{sent_nr}_{i,x}$, если $\text{sent_nr}_{i,x} \neq \emptyset$, иначе $-\infty$.

$\text{Min_rec_nr}_{i,x}$ = наименьшая временная метка в сообщениях в $\text{rec_nr}_{i,x}$, если $\text{rec_nr}_{i,x} \neq 0$, иначе $+\infty$.

Основываясь на свойствах временных меток Лампорта, можно доказать следующие свойства [2]:

Свойство 1. Если возможно пронумеровать все ЛКТ и сообщения в вычислении таким образом, чтобы для всех $c_{i,x}$ выполнялись следующие 2 условия:

$$C1_Z \equiv c_{i,x}.t < c_{i,x+1}.t$$

$C2_Z \equiv \max_rec_{i,x} \leq c_{i,x}.t \leq \min_sent_{i,x}$, то все ЛКТ принадлежат по крайней мере к одной консистентной ГКТ.

Свойство 2. Если возможно пронумеровать все ЛКТ и сообщения в инверсном вычислении таким образом, чтобы для всех $c_{i,x}$ выполнялись следующие 2 условия:

$$C1_{NR} \equiv c_{i,x}.t < c_{i,x+1}.t$$

$C2_{NR} \equiv \max_sent_nr_{i,x} \leq c_{i,x}.t \leq \min_rec_nr_{i,x}$, то пропущенных сообщений в распределенном вычислении нет.

Рассмотрим $C1_{ZNR} \equiv (C1_Z \wedge C1_{NR})$ и $C2_{ZNR} \equiv (C2_Z \wedge C2_{NR})$. Общий протокол удовлетворяет условиям $C1_{ZNR}$ и $C2_{ZNR}$ посредством:

- расставления временных меток для ЛКТ и сообщений;
- записи сообщений;
- направления процессов к взятию вынужденных ЛКТ.

Пусть lc_i будет временем по Лампорту для процесса P_i (lc_i дает $c_{i,x}.t$ свое значение, когда P_i берет $c_{i,x}$). Протокол описывается на абстрактном уровне, где в дополнение к lc_i используются 4 временные переменные и множество сообщений. Этими переменными являются MIN_SENT_i , MAX_REC_i , MAX_SENT_NR_i , MIN_REC_NR_i . MIN_... и MAX_... инициализированы $+\infty$ и $-\infty$ соответственно.

Взятие ЛКТ. Условие $C1_{ZNR}$ поддерживается процедурой checkpoint, в которой lc_i увеличивается и никогда не уменьшается другими процедурами. Данная процедура вызывается процессом P_i , когда он берет (по своей собственной инициативе) основную ЛКТ или когда он направляется протоколом для взятия вынужденной ЛКТ. Процедура checkpoint гарантирует, что время lc_i , используемое для отметок ЛКТ процесса P_i , всегда получает приращение.

Основная идея для выполнения второго условия заключается либо в направлении процесса к взятию вынужденной ЛКТ, либо в записи сообщений, когда условие может быть нарушено. Более точно, протокол гарантирует, что следующее условие всегда выполняется:

$$I \equiv \max(\max_rec_{i,x}, \max_sent_nr_{i,x}) \leq lc_i \leq \min(\min_sent_{i,x}, \min_rec_nr_{i,x}).$$

Посылка сообщения. Процедура send должна выполняться протоколом всякий раз, когда процесс P_i посылает сообщение m . Отметим, что когда сообщение m записано отправителем, отправитель информирует получателя, что m было записано (булево поле $m.recorded$), и это позволяет дважды не записывать одни и те же сообщения. В соответствии со значением, выбираемым для $m.t$, протокол может быть реализован несколькими способами. Процедура обеспечивает выполнение условия I по следующим причинам:

– выбор временной метки $m.t < lc_i$ не разрешен, так как он ведет к нарушению I ($MIN_SENT_i < lc_i$ и lc_i не может быть уменьшено);

– выбор временной метки $m.t = lc_i$ разрешен: I при этом не изменяется;

– выбор временной метки $m.t > lc_i$. В этом случае отправка m без его записи или без модификации значения lc_i приведет к тому, что $MAX_SENT_NR_i = m.t > lc_i$, нарушая тем самым левую часть I . Для предотвращения данного факта можно использовать 3 возможности:

(1) Взятие вынужденной ЛКТ. Это сбросит все переменные, появляющиеся в левой и правой части I .

(2) Запись сообщения m . В этом случае $m.t$ не будет принято во внимание в $MAX_SENT_NR_i$ и, соответственно, так как $m.t > lc_i$, I будет оставаться истинным. Это делает левую часть I нечувствительной к m .

(3) Изменение lc_i . Незапись сообщения m влечет за собой $MAX_SENT_NR_i = m.t > lc_i$. Таким образом, для истинности левой части I необходимо установить $lc_i = m.t$. Для сохранения истинности правой части I , необходимо ограничить данную возможность случаем, когда $m.t \leq \min(MIN_SENT_i, MIN_REC_NR_i)$.

Получение сообщения. Процедура *delivery* должна быть выполнена, когда сообщение поступает процессу P_i . Данная процедура обеспечивает поддержку I следующим образом:

– если $m.t < lc_i$: для обеспечения истинности правой части I сообщение m исключается из I посредством его записи (если оно не было записано прежде);

– если $m.t = lc_i$: I остается истинным;

– если $m.t > lc_i$: установление значений lc_i и MAX_REC_i в $m.t$ будет сохранять истинной левую часть I , таким образом в конце процедуры *delivery*: $MAX_REC_i = m.t = lc_i$.

В соответствии с правой частью I существуют 3 альтернативы:

(1) $m.t \leq \min(MIN_SENT_i, MIN_REC_NR_i)$. В этом случае правая часть I остается истинной.

(2) $m.t > MIN_SENT_i$. Для обеспечения условия $MAX_REC_i \leq MIN_SENT_i$ P_i направляется для взятия вынужденной ЛКТ.

(3) $MIN_REC_NR_i < m.t \leq MIN_SENT_i$: В этом случае $MIN_REC_NR_i$ должно быть установлено в значение $\geq m.t$. При этом существуют две возможности: либо направить процесс к взятию вынужденной ЛКТ, либо записать все сообщения $m' \in REC_NR_i$, такие что $m'.t < m.t$.

Важным фактом, вытекающим из этого протокола, является выбор между записью сообщений и взятием вынужденных ЛКТ. Это позволяет достичь компромисса между числом вынужденных ЛКТ и числом записанных сообщений.

Рассматривая описанные и предложенные подходы к определению согласованного глобального состояния распределенной системы вычислений можно отметить, что данная задача является отнюдь не тривиальной и требует серьезных теоретических исследований. Выбор для использования каждого подхода диктуется и определяется конкретной задачей и ее особенностями, а также предметной областью. На некоторые системы могут быть наложены строгие требования, как, например, отсутствие или малый объем постоянной па-

мяти, большой интервал между контрольными точками, что значительно сужает возможности выбора подходов. Коммуникационно-вынужденный подход является наиболее предпочтительным в современных системах, отличающихся своим масштабом и динамической масштабируемостью, а также значительной загрузкой каналов связи: в нем не требуется никаких дополнительных управляющих сообщений, он не имеет задержки при синхронизации и является прозрачным для конечного пользователя. Однако достигается это за счет дополнительного существенного усложнения используемых алгоритмов вычисления консистентных ГКТ, а также значительного увеличения объема сохраняемых данных. Дополнительную проблему создает задача максимально быстрой реакции на отказ какого-либо процесса, ведущая к увеличению требуемых ресурсов для реализации алгоритмов (как было показано, определение консистентного «среза» должно выполняться в реальном времени).

Встраивание предложенного подхода определения контрольных точек в общую модель восстановления работоспособности распределенной системы потребует дополнительного «окрашивания» машин, событий и сообщений метрическими данными и функциональной нагрузкой.

Рассматривая конкретные примеры реализации распределенных систем, в ряде случаев бывает недостаточно применения алгоритмов, позволяющих только зафиксировать отказы процессов, корректно их обработать и заново согласованно продолжить работу системы. Необходимо выявлять и анализировать причины возникновения таких отказов, что в конечном итоге может вылиться в изменение числа процессов, составляющих распределенную систему, и в сокращение или перераспределение возложенной на них функциональной нагрузки. Рассмотренный коммуникационно-вынужденный подход для определения глобального состояния системы является хорошей основой для проведения такого рода анализа и послужит фундаментом для создания систем с возможностью самовосстановления.

СПИСОК ЛИТЕРАТУРЫ

1. Chandy, K. M., Lamport, L., Distributed Snapshots: Determining Global States of Distributed Systems // ACM Transactions on Computer Systems. 1985. №3(1). P. 63-75.
2. Helary J-M., Mostefaoui A., Raynal M. Communication-induced Determination of Consistent Snapshots // IRISA Research Report. 1998. №0731-3071/98.
3. Elnozahy E. N., Johnson D. B., Wang Y. M. A Survey of Rollback-Recovery Protocols in Message-Passing Systems // ACM Computing Surveys. 1996.
4. Mostefaoui A., Raynal M. Efficient Message Logging for Uncoordinated Checkpointing Protocols // INRIA Rapport de recherche. 1996. №2972.

M. S. Koupriyanov, A. V. Kochetkov

METHODS OF CONSISTENT SYSTEM STATE DEFINITION

The basic methods of definition of a consistent state of distributed system, grounded on taking checkpoints are described. The communication-induced approach for taking checkpoints is offered and the procedures for send and receive messages are shown in given approach.

Consistent system state, local and global checkpoints, orphan messages, missing messages, communication-induced approach

ПИРАМИДАЛЬНО-РЕКУРСИВНАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

Рассматриваются вопросы пирамидального представления изображений и систематизация алгоритмов сжатия и восстановления изображений на основе рекурсивного разбиения. Представлены результаты моделирования одним из способов разбиения.

Цифровая обработка изображений, сжатие изображений, рекурсивные алгоритмы сжатия, регулярно-иерархические, пирамидальные, конусные структуры

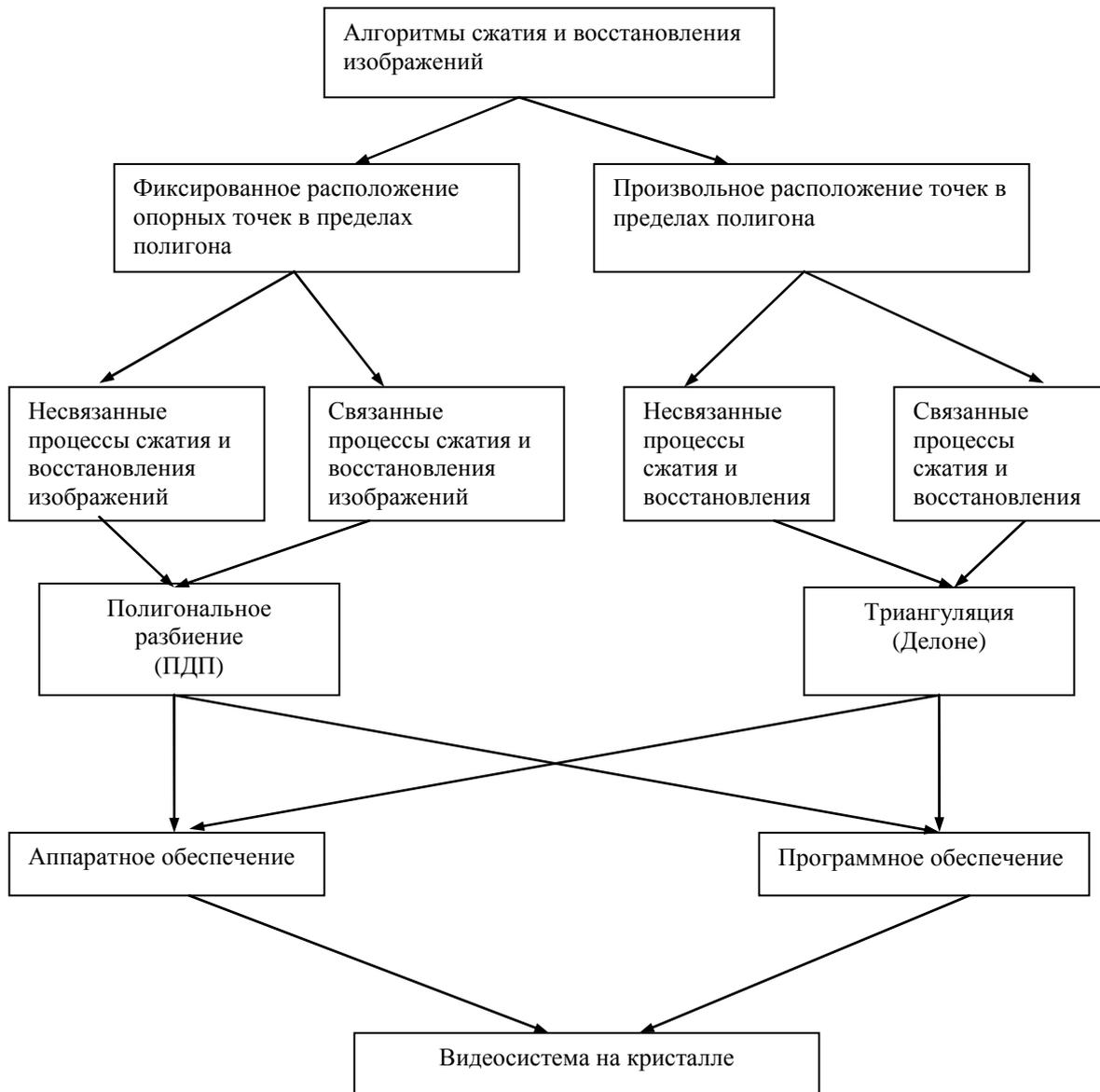
В настоящее время в области цифровой обработки изображений исследуются адаптивные методы обработки, которые наиболее полно учитывают свойства реальных изображений. В частности, перспективным направлением является использование рекурсивных структур данных для представления изображений в сжатом виде, что значительно снижает информационную избыточность последних [1], [2]. Целью статьи является рассмотрение и классификация наиболее распространенных пирамидально-рекурсивных методов при обработке и передаче неподвижных изображений.

Избыточность изображения обусловлена значительной корреляционной связью между яркостями смежных растровых элементов, которая проявляется в том, что основная площадь изображения занята полями, имеющими постоянную или мало изменяющуюся яркость, в то время как на мелкие детали и резкие световые переходы приходится лишь незначительная часть.

Актуальным является направление представления и обработки изображений с использованием рекурсивных (регулярно-иерархических, пирамидальных, конусных) структур. При представлении рекурсивной структурой изображение разбивается на области (полигоны) одинаковой яркости, а затем процедура разбиения рекурсивно повторяется для каждого полигона до тех пор, пока его размер не станет равным размеру наименьшего элемента исходного изображения. Такой процесс называется рекурсивной или регулярной декомпозицией. Каждому элементарному блоку приписывается значение, называемое «яркостью» или «цветом» и являющееся его интегральной характеристикой. По завершении процесса декомпозиции получаем набор изображений, последовательно уточняющих друг друга и сходящихся к исходному; этот набор обычно называют пирамидой изображения. Такие методы появились сравнительно недавно и получили название пирамидальных. Пирамидальные алгоритмы сжатия широко применяются в таких областях, как двумерное, векторное квантование, техническое зрение и при разработке теоретических основ, методов, алгоритмов, технологий и систем для обработки и распознавания изображений. В рамках этих направлений исследуются и решаются следующие проблемы и задачи:

- сегментация и распознавание полутоновых изображений;
- распознавание структуры и формы объектов;
- автоматизация поиска и анизотропических свойств изображений;
- представление изображений в базах данных и знаний.

На рисунке представлена обобщенная структурная схема реализации процессов рекурсивного сжатия изображений.



Регулярность пирамидальных структур предопределяет их удобную реализацию и эффективное использование в системах за счет параллельной обработки. Иерархия описаний различной степени общности способствует контекстно-независимой структуризации изображений. Эта структуризация служит основой для моделирования индуктивных (при вводе данных) и дедуктивных (при анализе пирамиды «сверху вниз») процессов человеческого восприятия. Действительно, если алгоритм представления видеоданных не зависит от их содержания и определен заранее, то имеется возможность быстрого параллельного ввода видеоданных в систему одновременно с получением обобщенного описания. Получение такого обобщения не требует выделения семантической структуры изображений, т. е. минует долгий аналитический путь восприятия [1], [2]. Анализ методов сжатия на основе пирамидально-рекурсивного подхода показал, что все они так или иначе преследуют цель сокращения избыточности на изображении, где восстановление производится присвоением областям одной яркости (например, средней), но отсутствуют способы выделения существенных областей (объектов) на исходном изображении. Рассмотренные особенности рекурсивных алго-

ритмов позволяют получить компактный программный код сжатия исходного изображения, что дает возможность выполнить аппаратную реализацию процессов рекурсивной декомпозиции. Особенно перспективным является использование при аппаратной реализации алгоритмов технологии «система на кристалле».

В данной статье предлагается систематизация, включающая объединение достоинств перечисленных направлений: развитие систем на кристалле, автоматизация и распараллеливание структур данных и пирамидально-рекурсивные методы при обработке и передаче неподвижных изображений.

Основные выводы:

1. Предлагается оценивать качество восстанавливаемых изображений в зависимости от сочетания как амплитудных (по яркости), так и геометрических (по координатам точек на плоскости) искажений при соблюдении баланса (эквивалентное влияние обоих факторов).

2. Найдено оптимальное сочетание амплитудных и геометрических отклонений при многократном проведении экспериментов в алгоритмах с фиксированным расположением опорных точек (ОТ) с целью достижения высокой степени сжатия и приемлемого качества.

3. В результате моделирования определены основные характеристики для аппаратной реализации алгоритмов (см. таблицу), положенных в основу оценки

| Вид оригинала | Коэффициент сжатия (S) | P_1 | P_2 | P_3 |
|---------------|------------------------|-------|-------|-------|
| Рельеф | 0,04 | 2412 | 1,39 | 326 |
| Портрет | 0,9 | 61306 | 1,25 | 8451 |
| Архитектура | 0,9 | 32612 | 1,66 | 4258 |

производительности, на базе технологии «система на кристалле»:

P_1 – длина сжатого описания в битах;

P_2 – плотность описания ОТ в битах;

P_3 – количество ОТ.

СПИСОК ЛИТЕРАТУРЫ

1. Прэйт У. Цифровая обработка изображений/ Пер. с англ. М.: Мир, 1982. Кн.1 и 2/
2. Фоли Дж., Дэм А. Основы интерактивной машинной графики/ Пер. с англ. М.: Мир, 1985. Кн.1 и 2.

A.I. Laristov, S.S. Fachmi

PYRAMIDAL - RECURSIVE PROCESSING OF IMAGES

Questions of pyramidal representation of images and ordering of algorithms of compression and restoration of images are considered(examined) on the basis of recursive splitting. Results of modelling are submitted to one of ways of splitting.

Digital processing of images, compression of images, recursive algorithms of compression, regular - hierarchical, pyramidal, cone structures

ВВЕДЕНИЕ В ПРОФЕССИЮ

Рассмотрена проблема повышения качества выпускников высшей школы посредством придания учебному процессу в высшей школе характера профессиональной производительной коллективнообразующей деятельности. Не затрагиваются вузы искусства, спортивные и военные, а также двухступенчатая схема «бакалавр – магистр».

Профессия, профессиональная высшая школа, придание учебному процессу в высшей школе характера профессиональной производительной коллективнообразующей деятельности

Попытаемся в кратком рабочем определении связать «профессию», «специальность», «квалификацию», чтобы все они были перед глазами.

Итак: «профессия – это избранное человеком вознаграждаемое ответственное квалифицированное создание специальной продукции (достигнутый результат, оказанная услуга), востребованной потребителем». Такой продукцией могут быть: теория, факт, мировой новизны способ, устройство, состав и др.; обобщая, в широком смысле, – польза заказчика, потребителя.

Кто такой «профессионал»? Человек, который объявляет эту профессию своей. А опыт работы? Профессиональная деятельность – ее содержание, продолжительность и результаты.

Далее, обратимся к высшей школе. Приходится признать тот факт, что вуз не является профессиональной высшей школой, вуз – это общеобразовательная высшая школа со специальным уклоном. Важным отличительным признаком этих образовательных концепций служит результат:

- результат «Профессия» плюс «Диплом», т. е., придание учебному процессу в вузе характера профессиональной производительной коллективнообразующей деятельности, – признак профессиональной высшей школы;

- результат «Диплом» без «Профессии» – признак общеобразовательной высшей школы со специальным уклоном.

Действительно, работодатель от соискателей вакансии требует «В.О. плюс О.Р. несколько лет», т. е. ему нужен профессионал, а у свежее испеченного выпускника, ищущего работу по дипломной специальности, опыта нет, есть только диплом, по сути – ученический. Понадобилось прекращение обязательного распределения и обязательного приема выпускника на постоянное место работы, чтобы давно очевидное было доказано всем: вуз не является профессиональной высшей школой. Все формы соединения обучения с производительным трудом не меняют школярской сути общеобразовательной высшей школы со специальным уклоном.

А то, что вуз служит общеобразовательной высшей школой со специальным уклоном, доказано многими студентами как во времена обязательного распределения на работу, так и в наши дни. Доказано очень просто: они поступают в вуз заведомо не ради профессии, но только ради статуса студента и диплома, причем получают желаемое; профессия им ни к чему.

Кто ущемлен тем, что вуз работает только в концепции общеобразовательной высшей школы со специальным уклоном, что вуз не работает в концепции профессиональной высшей школы? Это работодатели во всех формах собственности. Это студенты, настроенные вживаться в профессию, но вынужденные обучаться в общеобразовательном вузе со специальным уклоном. Это преподаватели, которые заинтересованы в профессиональном партнерстве со студентами. Это, наконец, сам вуз.

Идея взаимодействия в вузе обеих перечисленных образовательных концепций вызывает к жизни «Введение в профессию». «Введение в профессию» включает в себя:

- Процесс вживания студентов в профессию (определение профессии см. ранее) именно в вузе и с первого дня, т. е. осуществление концепции профессиональной высшей школы как альтернативы концепции общеобразовательной высшей школы со специальным уклоном.

- Рассказ «Введение в профессию». Этим словом – рассказ – усиленно подчеркивается отличие повествования от профессиональной деятельности в процессе вживания.

Рассказ «Введение в профессию» предназначен – как ориентирование – всем студентам, потому что:

- одни студенты, те, кто вживается в профессию, не избежат взаимодействия с ветвью вуза, работающей в концепции общеобразовательной высшей школы со специальным уклоном;

- другие студенты, которые первоначально были намерены вживаться в профессию, могут почему-то передумать, отказаться от вживания и удовлетвориться общеобразовательной высшей школой со специальным уклоном;

- третьи студенты, кому профессия была не нужна, могут изменить свои планы настолько, что решат вживаться в профессию;

- наконец, те студенты, которым эта профессия ни к чему, после окончания вуза будут – силою обстоятельств – принуждены избрать себе профессию, так что их тоже надо ориентировать.

Теперь коротко о рассказе «Введение в профессию».

Главное – студенту не надо ничего произвольно запоминать. Только ориентирование: опознавание, отличие, различение, сопоставление, соотнесение. Конечно, произвольное запоминание имеет место и ориентирование получается смешанным, т. е. с опорой на текст и по памяти. Студентам для ориентирования предлагается два сквозных задания:

- одно задание – по тексту: разметить, структурировать и переструктурировать текст;

- другое задание – по параметрам поисковой ситуации: выявить «фигуру» (на что направлено внимание), «фон» (что в данный момент второстепенно) и переформулировать ситуацию.

В рассказе «Введение в профессию» всего пять тем:

1. Ценности, смыслы, идеи, приоритеты.
2. Механизмы целеполагания.
3. Профессиональная деятельность.
4. Две концепции высшей школы (общеобразовательная высшая школа со специальным уклоном и профессиональная высшая школа).
5. Конкретная профессия, специализация, квалификация.

Несколько слов о каждой теме.

- Тема «Ценности, смыслы, идеи, приоритеты». Чего студент не хочет, чего он хочет, что ему на самом деле надо. Студент отличает скромность поведения от скромности профессиональных целей и результатов.

- Тема «Механизмы целеполагания». Студент различает цели, задачи, выдвижение и разложение цели. Результат как системообразующий элемент. Этапы решения задачи.

- Тема «Профессиональная деятельность». Студент различает профессию, специальность, квалификацию. Функции, обязанности, ответственность, средства, права, власть.

• Тема «Образовательные концепции». Студент отличает обе концепции – общеобразовательную и профессиональную – по конечному результату каждой и по ряду существенных характеристик, например: для чего, кого, чему и сколь долго, кому, как, за счет каких финансовых средств учить. Взаимодействие концепций и чем чреват для студентов переход из одной концепции в другую.

• Тема «Конкретная профессия, специальность, квалификация». Характеристика области деятельности, базовые науки, практика, подготовка кадров. Достижения, возможности, проблемы. Ориентирование студентов в квалификационных документах по дипломным проектам и дипломным работам конкретного вуза. Пособием для ориентирования служат методические рекомендации выпускающих кафедр этого вуза. Студенты реферируют дипломные проекты или дипломные работы, и этот реферат является отчетным документом.

Требование к языку рассказа «Введение в профессию»: понятность студентам.

Полагаем, что поднятые проблемы и предложения по их решению актуальны, способствуют развитию высшей школы и повышению качества ее выпускников.

A.V.Kuupper

DRAWING INTO PROFESSION

The article is dedicated to the problem of the improvement the quality of the graduates by means of the imparting the educational process in the higher school the character of professional productive collectiveforming activity. The article doesn't touch upon art, sport, and military higher educational establishments.

Profession, professional higher school, problem of the improvement the quality of the graduates by means of imparting the educational process in higher school the character of professional productive collectiveforming activity

УДК 621.3.049.77.001.2

А. Б. Чирков

ПРИМЕНЕНИЕ САМООРГАНИЗУЮЩЕЙСЯ КАРТЫ КОХОНЕНА ДЛЯ РЕШЕНИЯ ЗАДАЧИ КЛАСТЕРИЗАЦИИ

Описывается применение самоорганизующейся карты Кохонена для решения задачи кластеризации. Описана процедура обучения без учителя с использованием модифицированного правила обучения Хебба (алгоритм Self-Organizing feature Map, SOM), рассмотрено применение алгоритма SOM как алгоритма векторного квантования.

Самоорганизация, карты Кохонена, нейронные сети, правило обучения Хебба, кластеризация, векторное квантование, обучение без учителя

Самоорганизующиеся нейронные сети Кохонена (*карты Кохонена*) возникли в результате исследований зрительной области коры головного мозга [2, 3]. В ходе этих исследований выяснилось, что ряд областей мозга (как правило, это области, ответственные за первоначальную обработку сенсорной информации) представляют собой топологически упорядоченные структуры (так называемые *топографические карты*), которые обладают следующими свойствами [1]:

• Упорядоченность отображения входов сенсоров в соответствующие физические области этих структур.

• Нейроны, которые обрабатывают похожие входные данные, физически расположены близко друг к другу и способны влиять на соседей посредством синаптических связей.

- Между ближайшими нейронами-соседями возникает положительная обратная связь (т. е. возбужденный нейрон способен частично активировать близлежащие нейроны), а между отдаленными – отрицательная обратная связь (так называемый эффект латерального торможения).

Исходя из этих свойств, можно сделать следующий вывод: пространственное положение выходных нейронов в топографической карте соответствует конкретной области признаков данных, выделенных из входного пространства, т. е. в данных структурах происходит предварительная кластеризация входных данных по признакам. На рис. 1 приведен фрагмент структуры топографической карты, где серым кружком обозначен вход сети, прямыми линиями – взвешенные синаптические связи входа с нейронами карты, светлыми кружками – неактивированные нейроны, а черным кружком – активированный нейрон-победитель.

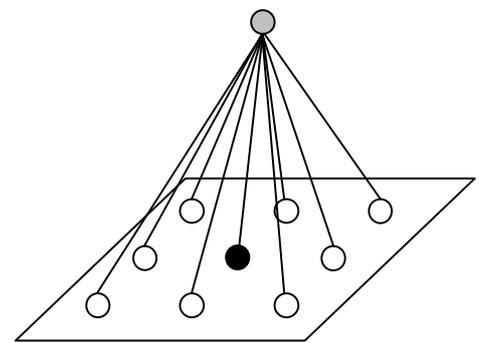


Рис. 1

Из приведенных свойств реальной биологической топографической карты можно представить ее эвристическую математическую модель. В работе Кохонена [4] такая математическая модель топографической карты получила название самоорганизующейся карты признаков (Self-Organizing feature Map, SOM). Основной задачей самоорганизующейся карты является преобразование непрерывного многомерного пространства входных признаков в дискретное пространство, представленное решеткой нейронов. Размерность решетки принимается обычно равной 2 или 1. Синаптические связи обеспечивают взвешенную связь всех входов с каждым нейроном карты. На рис. 2 приведена структура двумерной решетки размера $M \times K$.

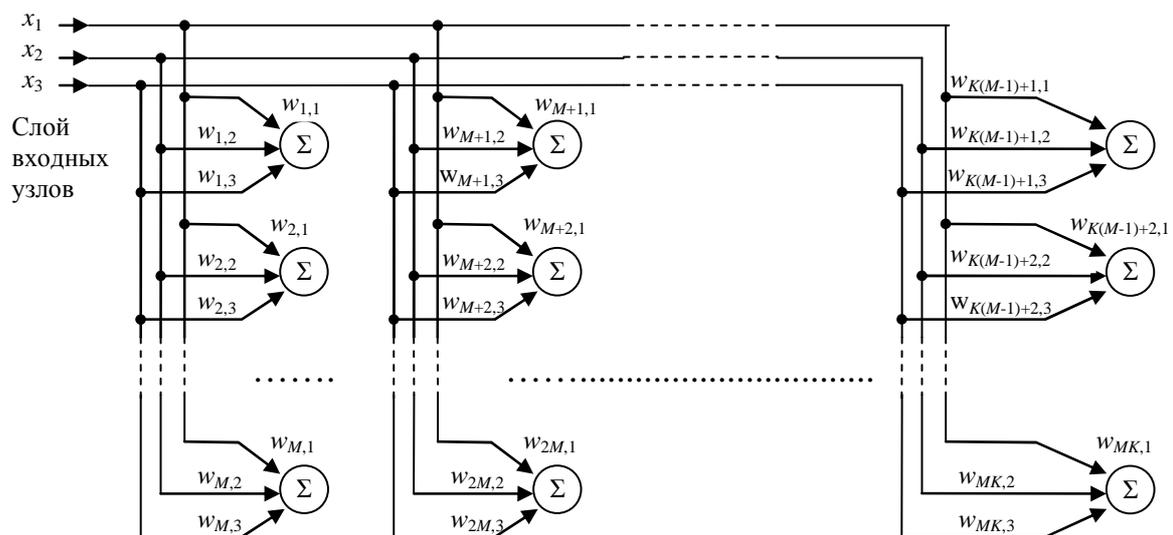


Рис. 2

Самоорганизующаяся карта Кохонена представляет собой нейронную сеть, базирующуюся на конкурентном методе обучения без учителя. Обучение без учителя означает, что в процессе обучения нейронной сети ей не предъявляются пары из входного вектора и желаемого выходного вектора, а обучение происходит по так называемому *принципу обучения Хебба*.

Принцип обучения Хебба основан на гипотезе изменения веса синаптических связей между активными нейронами [1]. Математически изменение веса Δw_{ij} синаптической связи между j -м источником сигнала и i -м нейроном на шаге обучения n можно выразить как функцию входного сигнала синапса и выхода нейрона:

$$\Delta w_{ij} = F(y_i(n), x_j(n)), \quad (1)$$

где $x_j(n)$ – j -й элемент входного вектора $\vec{x}(n)$, поступающий на вход синапса j , $y_i(n)$ – выходной сигнал i -го нейрона. В векторной форме записи простейший вид правила обучения Хебба (1) выглядит следующим образом:

$$\Delta \vec{w}_i = \eta y_i(n) \vec{x}(n), \quad (2)$$

где η – коэффициент *скорости обучения*, который обычно уменьшается в процессе обучения. Однако в случае конкурентного обучения данный вид правила Хебба оказывается неприемлемым, поскольку в процессе обучения веса изменяются только в сторону увеличения, что рано или поздно приводит все веса сети в состояние насыщения. Введем в формулу (2) *слагаемое забывания* вида $-g(y_i(n))\vec{w}_i(n)$, причем функция $g(y(n))$ неотрицательная, и $g(y_i(n)) = 0$ при $y_i(n) = 0$. Тогда правило обучения Хебба примет вид:

$$\Delta \vec{w}_i = \eta y_i(n) \vec{x} - g(y_i(n)) \vec{w}_i(n). \quad (3)$$

Пусть $g(y_i(n))$ имеет вид простейшей линейной функции: $g(y_i(n)) = \eta y_i(n)$, тогда выражение (3) примет вид:

$$\Delta \vec{w}_i = \eta y_i(n) (\vec{x} - \vec{w}_i(n)). \quad (4)$$

Конкурентный метод обучения подразумевает, что в процессе обучения нейроны сети соревнуются между собой за право активации, после чего веса синапсов нейрона-победителя, а также нейронов, находящихся в некоторой топологической окрестности нейрона-победителя, модифицируются в соответствии с правилом обучения Хебба таким образом, чтобы приблизить вектор синаптических весов к входному вектору. Процесс конкуренции основан на поиске нейрона с максимальным значением скалярного произведения $\vec{w}_i^T \vec{x}$, что эквивалентно поиску нейрона с вектором синаптических весов, максимально приближенным к входному вектору по Евклиду:

$$i(\vec{x}) = \arg \left(\min_j \|\vec{x} - \vec{w}_j\| \right). \quad (5)$$

Для характеристики топологической окрестности введем *функцию топологической окрестности нейрона в решетке* $h_{j,i}(d_{ij})$. Эта функция может определяться различными способами, но все они учитывают расположение нейрона относительно нейрона-победителя на топологической карте. Примером такой функции служит функция Гаусса:

$$h_{j,i}(d_{ij}) = \exp \left(-\frac{d_{j,i}^2}{2\sigma^2} \right),$$

где d_{ij} – латеральное расстояние между i -м и j -м нейронами, определяемое как расстояние в некоторой геометрической окрестности нейрона. Форма геометрической окрестности может быть квадратом, окружностью и т. д. (рис. 3). Параметр σ называют *эффективной шириной*, и он характеризует дисперсию функции топологической окрестности. В классическом алгоритме SOM эффективная ширина является, в общем случае, непостоянной величиной, уменьшающейся с течением времени.

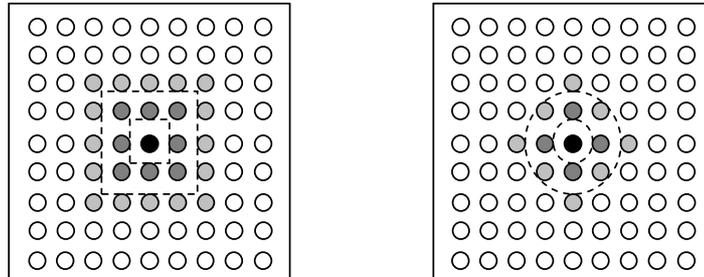


Рис. 3

Если принять значение выхода нейрона-победителя d формуле (4) за максимальное значение функции топологической окрестности, то формула (4) примет вид:

$$\Delta \bar{w}_j = \eta h_{j,i}(d_{i,j})(\bar{x} - \bar{w}_j(n)) .$$

Итак, выделим основные шаги обучения карты Кохонена:

1. Инициализация векторов синаптических весов $\bar{w}_i(n)$ небольшими случайными значениями.

2. Выборка входного вектора \bar{x} из пространства входных векторов.

3. Поиск нейрона с максимально приближенным к входному вектору \bar{x} вектором синаптических весов (процесс конкуренции) по формуле (5). Степень близости векторов \bar{x} и $\bar{w}_i(n)$ обычно определяется как Евклидово расстояние : $r_i^2 = \|\bar{x} - \bar{w}_i\|$. Нейрон с минимальным расстоянием r_i объявляется победителем.

4. Коррекция синаптических весов нейронов по формуле :

$$\bar{w}_j(n+1) = \bar{w}_j(n) + \eta(n)h_{j,i}(d_{i,j})(\bar{x} - \bar{w}_j(n)) ;$$

где i – номер нейрона-победителя; $\bar{w}_j(n)$ – текущий вектор синаптических весов нейрона j ;

$\eta(n)$ – текущее значение коэффициента скорости обучения ; $h_{j,i}(d_{i,j})$ – функция окрестности с центром в победившем нейроне i . Скорость обучения и эффективная ширина функции окрестности $\sigma(n)$ в общем случае меняются в процессе обучения для получения наилучшего результата. Так, в [1] предлагается вычислять параметры обучения следующим образом:

$$\eta(n) = \eta_0 \exp\left(-\frac{n}{\tau_1}\right) ;$$

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_2}\right) ;$$

где τ_1, τ_2 – параметры, задаваемые перед началом обучения. Однако в ряде практических применений достаточно задать фиксированное значение σ и значение η в пределах 0.1 ... 0.01 на различных этапах работы алгоритма.

1. Возврат к шагу 2 до тех пор, пока в карте признаков не перестанут происходить заметные изменения.

В процессе обучения карты Кохонена можно выделить 2 этапа: *этап самоорганизации* и *этап сходимости* [1]. На этапе самоорганизации происходит топологическое упорядочение векторов весов. Этот этап может занять около 1000 итераций алгоритма, при этом коэффициент скорости обучения должен иметь значение около 0.1, а функция топологической окрестности должна охватывать все нейроны сети, однако по завершении этого этапа функция окрестности должна охватывать только ближайших соседей нейрона-победителя.

На этапе сходимости происходит подстройка весовых векторов нейронов для точного квантования входного пространства. На этом этапе коэффициент скорости обучения должен быть примерно 0.01, а весь этап занимает порядка $500 * N$ итераций, где N – число нейронов сети.

После завершения процесса обучения карта Кохонена представляет собой *карту признаков*, которая отображает непрерывное входное пространство X в *дискретное пространство признаков* A , т. е. выполняет нелинейное преобразование: $\Phi : X \rightarrow A$.

Карта признаков обладает следующими свойствами [1]:

1. Карта признаков Φ , представленная множеством векторов синаптических весов $\{\bar{w}_j\}$ в выходном пространстве A , реализует хорошую аппроксимацию входного пространства X .

2. Карта признаков, полученная алгоритмом SOM, является топологически упорядоченной в том смысле, что положение нейронов в решетке соответствует конкретной области или признаку входного образа.

3. Карта признаков отражает вариации в статистиках распределения входного сигнала, т. е. области входного пространства X , из которого с большей вероятностью берутся входные векторы, занимают большие области в выходном дискретном пространстве A .

4. Для данных из входного пространства с нелинейным распределением карта признаков способна извлечь набор наилучших признаков.

В [1] показано, что алгоритм обучения SOM является по своей сути *алгоритмом векторного квантования*, работающим в последовательном режиме. Последовательный режим работы означает, что разбиение входного пространства на дискретное пространство признаков осуществляется по мере поступления данных, т. е. в режиме on-line. Также существует *пакетная* версия алгоритма, которая работает с выборкой входных данных и выделяет признаки из входного пространства в режиме off-line.

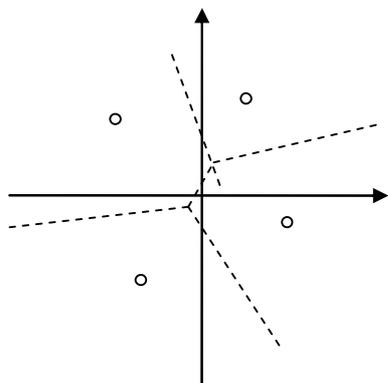


Рис. 4

Алгоритмы векторного квантования представляют собой способ сжатия данных за счет выделения существенных признаков из распределения входных векторов. Они разбивают входное пространство на так называемые *ячейки Вороного*, которые представляют собой области пространства, точки которых расположены ближе всего к ближайшему *вектору воспроизведения* [5]. Пример разбиения входного пространства на ячейки Вороного приведен на рис. 4.

Алгоритм SOM представляет собой приближенный метод вычисления векторов произведения без учителя, при этом аппроксимация векторов определяется векторами си-наптических весов нейронов на карте признаков. Очевидно, что обученная карта Кохонена, работающая по принципу конкуренции, является системой векторного квантования и решает задачу кластеризации входных данных по признакам.

СПИСОК ЛИТЕРАТУРЫ

1. Хайкин С, Нейронные сети: полный курс. 2-е изд. : Пер.с англ. // М.: Издательский дом «Вильямс», 2006.
2. Knudsen E.I., S. duLac and Esterly S.D. Computational maps in the brain. // Annual Review of Neuris- science. 1987. Vol.10. P. 41-65 .
3. Hubel D.H. , T.N. Wiesel. Functional architecture of macaque visual cortex. // Proc. of the Royal Society, London. 1977. Vol.198. P. 1-59 .
4. Kohonen T. Self-organized formation of topologically correct feature maps. // Biological Cybernetics. 1982. Vol.43. P.59-69 .
5. Gersho A., Gray R.M. Vector Quantization and Signal Compression. // Kluwer. Norwell. 1992.

A. B. Chirkov

APPLICATION KOHONEN'S SELF-ORGANIZING MAPS FOR THE DECISION OF A CLUSTERING PROBLEM

Application Kohonen's Self-Organizing maps for the decision of a problem clustering is described. Procedure of training without the teacher with use of the Hebb's learning rule of training is described (algorithm Self-Organizing feature Map, SOM), application of algorithm SOM as algorithm of vector quantization is considered.

Self-Organizing, Kohonen's maps, neural nets, Hebb's learning rule, clustering, vector quantization, learning without teacher

УДК 519.6

А. Н. Колодяжный

ПРИМЕНЕНИЕ ВЕЙВЛЕТОВ ДЛЯ ВЫЯВЛЕНИЯ АНОМАЛЬНЫХ ВЫБРОСОВ В ТРАЕКТОРНЫХ ИЗМЕРЕНИЯХ

Показано, что использование аппарата вейвлет-анализа для определения аномальных выбросов в измерениях (в частности, траекторных) более перспективно по сравнению с классическими методами, так как позволяет определить выбросы и в тех случаях, когда классические методы не работоспособны. Приводится алгоритм определения опорной выборки при наличии аномальных выбросов.

Вейвлет-преобразование, опорная выборка, аномальные измерения

Практика статистической обработки радиотехнических траекторных измерений показывает, что среди множества результатов измерений имеются отдельные значения, которые резко отличаются от всех остальных. Такие измерения называют аномальными или сингулярными [1]. Появление аномальных результатов измерений в данных ЭВМ, ведущей статистическую обработку измерений, обусловлено сбоями в работе измерительных средств, машинными сбоями при предварительной математической обработке измерений, сбоями при передаче данных по линиям связи. Возникновение сбоев определяется внешними и внутренними фак-

торами. Влияние внешних факторов (колебание питающих напряжений, воздействие вибрационных и ударных нагрузок, изменение температуры окружающей среды и т. д.) в значительной степени может быть исключено применением стабилизирующих устройств. Внутренние факторы (флуктуации параметров радиоэлементов, постепенное изменение граничных значений параметров радиоэлементов за счет их старения и т. п.) в основном характеризуются постепенными отказами, которые проявляют себя через сбои.

Аномальные измерения необходимо исключить из последующей обработки, так как они нарушают статистический характер информации изучаемого процесса, что приводит к искажению результатов при определении опытной траектории и ее характеристик, получаемых по результатам измерений. Способы отбраковки измерений базируются на априорных данных о характере изменения измеряемых функций во времени. Вместе с тем известно, что при устойчивых условиях наблюдений возможны весьма большие отклонения результатов отдельных измерений от их истинных значений, хотя вероятность таких отклонений очень мала. Это свойство измерений исключает возможность достоверного решения вопроса о принадлежности данного измерения, резко отличающегося от других, к имеющейся выборке. Теория вероятностей может дать только вероятностную оценку аномальности данного измерения.

Аномальные измерения исключают в процессе предварительной статистической обработки измерений, при этом классическим подходом может быть применение, так называемой, опорной выборки измерений[1]. При использовании данного метода результаты обработки могут оказаться не приемлемыми, так как формируемая опорная выборка может содержать более одного аномального измерения. В связи с этим стандартный алгоритм не позволяет сформировать опорную выборку и в итоговую совокупность измерений приходится включать аномальные измерения или исключать все, включая и хорошие. Использование вейвлет-технологий позволяет сформировать опорную выборку и при наличии в ней более одного аномального измерения, причем нет необходимости использовать подход, основанный на вейвлет-анализе всей выборки измерений, что заметно сокращает время обработки.

Вейвлет-преобразование. На основании понятия о векторном пространстве общепринятым подходом к анализу сигнала $s(t)$ стало его представление в виде линейной комбинации простых составляющих – базисных функций $\Psi_{jk}(t)$, помноженных на коэффициенты C_{jk} , с уровнем разложения j :

$$s(t) = \sum_j \sum_k C_{j,k} \Psi_{j,k}(t). \quad (1)$$

Так как базисные функции $\Psi_k(t)$ предполагаются заданными как функции вполне определенного вида, то только коэффициенты C_k содержат информацию о конкретном сигнале. Классические базисные функции, например многочлены, имеют один серьезный и принципиальный недостаток – они не способны адаптироваться к локальным изменениям сигналов. Вейвлеты решают данную задачу и образуют некоторый набор функций, удовлетворяющих сформулированным далее условиям, основанным на использовании представления сигнала в виде (1).

Термин «вейвлет» в переводе с английского означает «короткая или маленькая волна». Прямое вейвлет-преобразование, именуемое также непрерывным преобразованием, означает разложение произвольного входного сигнала по принципиально новому базису в виде совокупности волновых пакетов. Довольно грубо можно представить вейвлеты как некоторые волновые функции, способные осуществлять преобразование Фурье не по всей временной оси, а локально по месту своей локализации.

К сожалению, почти все вейвлеты не имеют аналитического представления в виде формулы, но могут задаваться итерационными выражениями, легко вычисляемыми компьютерами.

Вейвлеты характеризуются четырьмя принципиально важными свойствами:

- имеют вид коротких, локализованных во времени (или в пространстве) волновых пакетов с нулевым значением интеграла;
- обладают возможностью сдвига по времени;
- способны к масштабированию (сжатию/растяжению);
- имеют ограниченный (или локальный) частотный спектр.

Вейвлет-базис может быть ортогональным, что заметно облегчает анализ, дает возможность реконструкции сигналов и позволяет реализовать алгоритмы быстрых вейвлет-преобразований [2].

В основе непрерывного вейвлет-преобразования лежит использование двух непрерывных и интегрируемых по всей оси t (или x) функций [3]:

1) вейвлет-функции $\psi(t)$ с нулевым значением интеграла, определяющей детали сигнала и порождающей детализирующие коэффициенты;

2) масштабирующей или скейлинг-функции $\varphi(t)$ с единичным значением интеграла, определяющей грубое приближение (аппроксимацию) сигнала и порождающей коэффициенты аппроксимации.

Прямое непрерывное вейвлет-преобразование сигнала $s(t)$ задается по формальной аналогии с преобразованием Фурье путем вычисления вейвлет-коэффициентов по формуле:

$$C(a, b) = \int_{-\infty}^{\infty} s(t) a^{-1/2} \psi((t-b)/a) dt. \quad (2)$$

Прямое вейвлет-преобразование можно рассматривать как разложение сигнала по всем возможным сдвигам и растяжениям/сжатиям сигнала $s(t)$ или некоторой произвольной функции. При этом параметры a и b могут принимать любые значения в пределах областей их определения.

Исключение аномальных измерений. Существует большое количество вейвлетов, форма которых близка к виду некоторых сингулярных выбросов, поэтому такие вейвлеты неплохо описывают рассматриваемый сигнал, позволяя наиболее наглядно выявить все его локальные особенности. К таким вейвлетам можно отнести ортогональные вейвлеты с компактным носителем, например Добеши (рис. 1, а), Симлета (рис. 1, б), Койфлета (рис. 1, в), [4], а также биортогональные парные вейвлеты с компактным носителем (рис. 1, г). На графиках по горизонтальной оси отложено значение аргумента вейвлет-функции (ось времени), по вертикальной – значение вейвлет-функции.

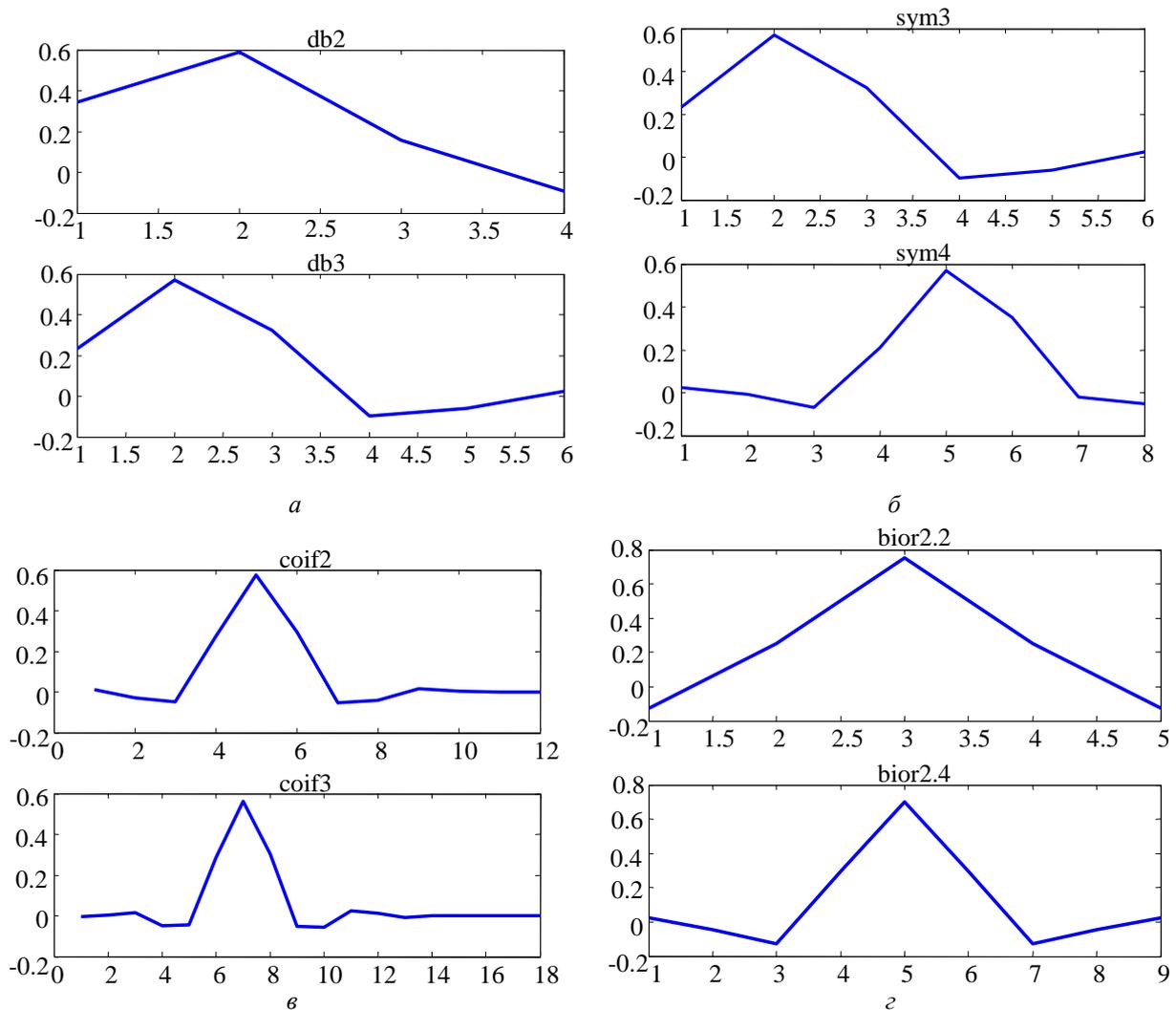


Рис. 1

Плюсом данных вейвлетов является также возможность применения к ним дискретного преобразования с использованием алгоритма быстрого вейвлет-преобразования [3], что уменьшает время вычислений.

Вейвлет-оценка сигнала, обладающего некоторыми выбросами, показывает наличие в данных высокочастотных составляющих. Нетрудно догадаться, что подобные составляющие могут быть устранены с помощью некоторой процедуры обработки коэффициентов детализации C_k , содержащих информацию о высокочастотной части спектра данных (рис. 2). Пример сигнала с аномальными измерениями и его вейвлет-спектрограмма представлены на рис. 2. Светлые полосы спектра соответствуют большим значениям коэффициентов детализации сигнала. Говоря конкретнее, коэффициенты детализации, обладающие «большими значениями», должны быть уничтожены введением некоторого порога. Процедура обнуления или пересчета коэффициентов детализации, значения которых превышают порог, представляет собой процедуру пороговой обработки, или «трешолдинг»[4].

Существует ряд процедур пороговой обработки коэффициентов детализации. Для решения задачи определения аномальных выбросов может быть использована процедура жесткого трешолдинга, предложенная Д. Донохо и Я. Джонстоном. Жесткий трешолдинг подразумевает категоричное обращение с коэффициентами детализации:

$$C_{j,k} = C_{j,k}(|C_{j,k}| < \tau), \quad (3)$$

где $\tau > 0$ – некоторое пороговое значение.

В качестве τ можно брать[4]:

$$\tau = \tau_{j,k} = \sqrt{2\sigma_{j,k}^2 \log M_j}, \quad (4)$$

где $\sigma_{j,k}$ – эмпирическое СКО ошибок измерений; M_j – число ненулевых коэффициентов разложения на каждом из уровней анализа.

Коэффициенты $C_{j,k}$ желательно нормировать:

$$C_{j,k} = C_{j,k} / \max_k \{C_{j,k}\}. \quad (5)$$

Экспериментально установлено, что пороговое значение τ для нормированных коэффициентов $C_{j,k}$ должно принадлежать интервалу (0.1; 0.3). Для случая определения сингулярных выбросов измерений достаточно рассмотреть коэффициенты первого уровня детализации $j = 1$.

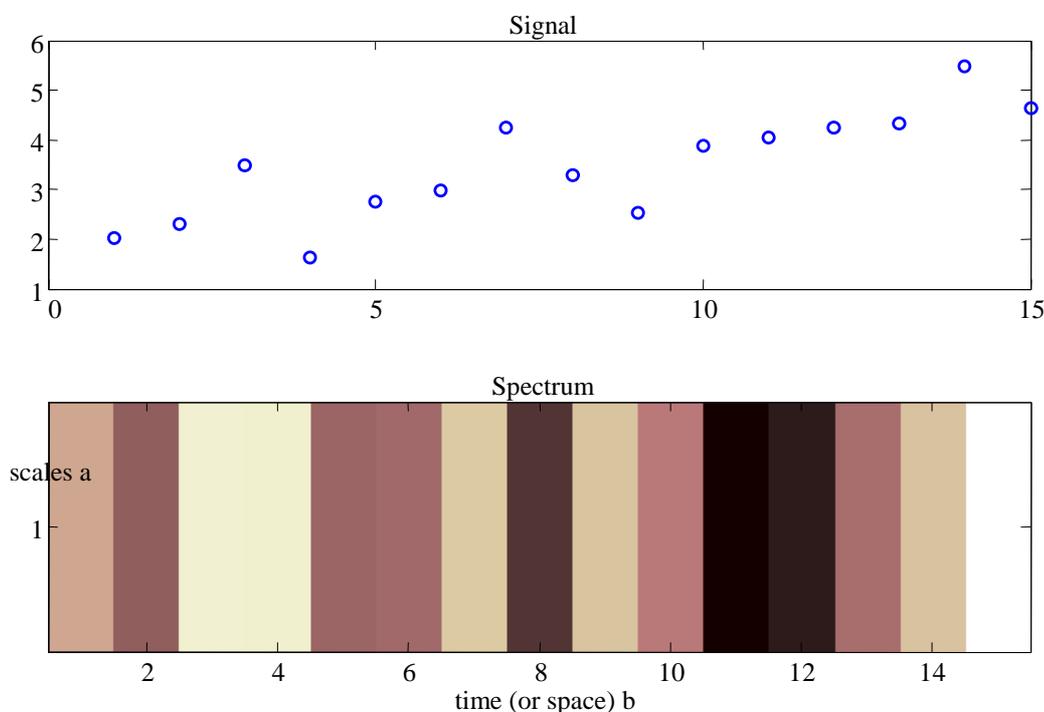


Рис. 2

При разложении исходного сигнала вейвлет коэффициенты для первого и последнего значения сигнала могут превосходить остальные коэффициенты разложения, не будучи аномальными. Это связано с особенностями используемых вейвлет-функций, поэтому для таких значений необходимо использовать другие критерии оценки аномальности.

Подытоживая все изложенное ранее можно сформулировать алгоритм определения опорной выборки измерений на основе вейвлет-технологий:

1. Составляют выборку из первых полученных ν измерений.
2. По формуле (4) вычисляют пороговое значение τ приняв $M = \nu$.
3. Рассчитывают коэффициенты C_k ($k = 1, \dots, \nu$) вейвлет-разложения и нормируют их по формуле (5).

4. Проверяют выполнение условия (3). Если неравенство верно, то k -е измерение включают в опорную выборку.

5. Используя соотношения (1)–(4)[1, с.145] проверяют, будет ли сформированная выборка опорной, т. е. отсутствуют ли в ней аномальные измерения. Если нет, то из выборки последовательно исключают 1-е значение, проверяя, будет ли сформированная выборка опорной, затем последнее, выполняя эту же проверку.

6. Если объем сформированной опорной выборки будет меньше заданного ν , то выборку дополняют следующими измерениями $k = \nu + 1, \dots$, проверяя по формулам (1)–(4)[1, с.145] не будет ли новое измерение аномальным. Если нет, то измерение включают в выборку. Если количество аномальных измерений, следующих друг за другом, превзойдет некоторое наперед заданное число m , то переходят к поиску новой опорной выборки.

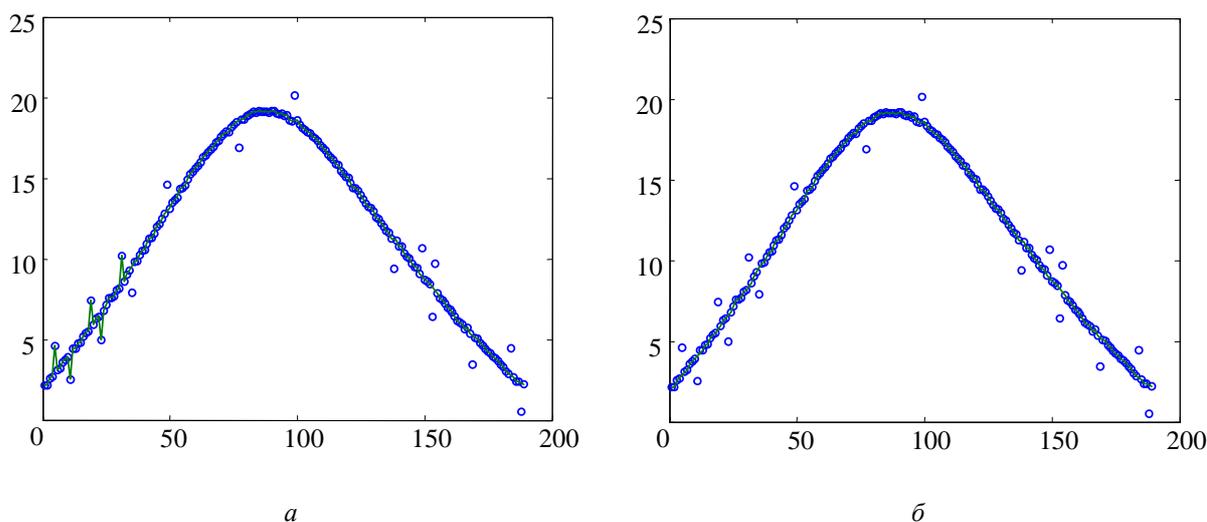


Рис. 3

Результаты. Разработанный алгоритм был реализован с использованием пакета Matlab6. Исходная измерительная последовательность зашумлялась нормальным шумом с заданным СКО σ , в случайные моменты были внесены аномальные измерения $k \cdot \sigma$ ($k = 3$), количество которых не превышало 10 % от общего количества измерений. Количество проведенных экспериментов – 100. В результате суммарное СКО сглаживаемого сигнала относительно истинной модели для алгоритма с использованием вейвлет-технологии оказалось вдвое меньше, чем при классическом подходе, при одном и том же порядке опорной выборки. Один из примеров обработки сигнала приведен на рис. 3, где точками обозначены исходные измерения, линией соединяются измерения, входящие в итоговую совокупность. Видно, что для одних и тех же исходных данных при использовании классического подхода (рис. 3, а) аномальные измерения не отбраковываются (выбросы в начале последовательности измерений), а при использовании вейвлетов (рис. 3, б) они локализируются и уже не входят в итоговый набор измерений.

Таким образом, результаты статистического моделирования подтверждают высокую эффективность предложенного алгоритма отбраковки аномальных измерений.

СПИСОК ЛИТЕРАТУРЫ

1. Жданюк Б. Ф. Основы статистической обработки траекторных измерений, М.: Советское радио, 1978
2. Дьяконов В. П. Вейвлеты. От теории к практике, М.: Солон-Р, 2002
3. Поликар Р. Введение в вейвлет-преобразование / перевод с английского В. Г. Грибунина,
4. Пикар Д. Вейвлеты, аппроксимация и статистические приложения / перевод с английского К. А. Алексева

It is shown, that use of the device of the wavelet – analysis for definition of abnormal emissions in measurements (in particular trajectories) is most perspective in comparison with classical methods as allows to define emissions and when classical methods are not efficient. The algorithm of definition of abnormal measurements is resulted.

Wavelet – transformation, basic sample, abnormal measurements

УДК 681.3.066

А. Н. Ветров, Е. Е. Котова, Н. Н. Кузьмин

АДАПТИВНАЯ ИНФОРМАЦИОННО-ОБРАЗОВАТЕЛЬНАЯ СРЕДА АВТОМАТИЗИРОВАННОГО (ДИСТАНЦИОННОГО) ОБУЧЕНИЯ НА ОСНОВЕ ПАРАМЕТРИЧЕСКИХ КОГНИТИВНЫХ МОДЕЛЕЙ

Предлагается подход к созданию, анализу и повышению эффективности функционирования информационно-образовательной среды автоматизированного (дистанционного) обучения на основе параметрических когнитивных моделей

Информационно-образовательная среда, система автоматизированного (дистанционного) обучения, когнитивная модель, блок параметрических когнитивных моделей, технология когнитивного моделирования, методика формирования структуры когнитивной модели.

Информатизация образовательных учреждений инициирует рассмотрение широкого спектра научных аспектов: региональный, экономический, организационный, технический, внедренческий, педагогический, методический, эргономический, психологический, физиологический, лингвистический.

Генезис идей автоматизированного обучения прослеживается с середины 20-х годов прошлого века и связан с появлением классических моделей используемых в основе образовательных средств: линейная (Б. Ф. Скиннер), линейная с обратной связью (С. Л. Пресси), разветвленная и разветвленная многоуровневая (Н. Краудер, Г. Паск), которые в дальнейшем получили развитие в работах ряда ученых: А. И. Берг, В. П. Беспалько, П. Я. Гальперин, В. М. Глушков, А. М. Довгялло, Л. Б. Ительсон, Л. Н. Ланда, А. Н. Леонтьев, Е. И. Машбиц, Н. Ф. Талызина, Е. Л. Ющенко и др.

Информационно-образовательные среды (ИОС) систем автоматизированного (дистанционного) обучения (АДО) созданные на основе традиционных моделей и технологий практически не учитывают индивидуальные особенности (характеристики) личности субъекта обучения (ЛХО). Исключением является разветвленная многоуровневая модель, учитывающая уровень остаточных знаний обучаемого и уровень изложения по изучаемому предмету.

Современный уровень развития науки и технологии позволяет реализовать средства обучения нового поколения учитывающие ЛХО на основе гибридной и адаптивной моделей, что существенно влияет на качество (пере)подготовки специалистов.

Индивидуальная ориентация информационного взаимодействия между субъектами и средствами обучения в ИОС АДО достигается за счет использования ряда технологий: технология индивидуального обучения – реализует схему «субъект – средство обучения»; технология индивидуализированного обучения – обеспечивает учет индивидуальных осо-

бенностей личности; технология адаптивного обучения – позволяет генерировать образовательные воздействия на основе инвариантной модели.

Для реализации адаптивной генерации информационно-образовательных воздействий согласно ЛХО в ИОС АДО предлагается использовать когнитивную модель (КМ), структура которой отвечает ряду требований: релевантность – сформировать модель с набором личностных характеристик субъекта учитываемых средствами ИОС и существенными для достижения целей обучения; адекватность – соответствие полученной модели оригиналу; состоятельность – поддержка средствами ИОС квазидинамического анализа и обновления параметров модели за счет систематического накопления данных о состоянии субъекта.

Структура обучения как технологического процесса формирования знаний обучаемого (рис. 1), включает ряд заделов (этапов): планирование, подготовка учебного-методического комплекса (УМК), фаза обучения, анализ и контроль.

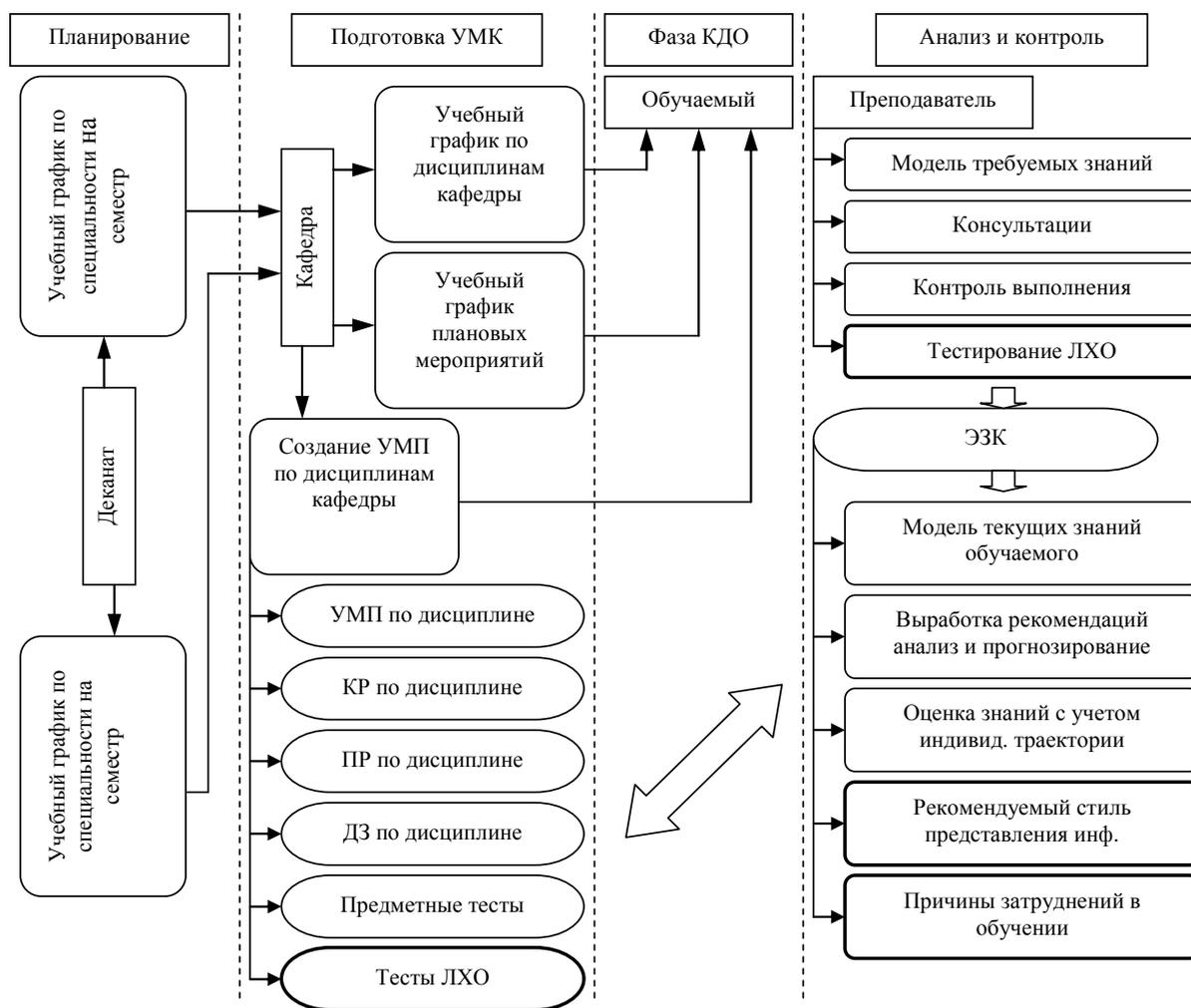


Рис. 1.

Технологические особенности при организации автоматизированного (дистанционного) обучения на основе параметрических когнитивных моделей

Реализация ИОС адаптивного обучения на основе блока КМ влечет модификацию некоторых этапов образовательного процесса: этап подготовки УМК – необходимо подготовить тесты ЛХО; этап анализа и контроля – требуется протестировать ЛХО, выявить предпочтительные параметры представления информационных фрагментов, причины затруднений обучаемого при их восприятии, а затем занести результаты в электронную записную книжку (ЭЗК).

Структура системы АДО с элементами адаптации на основе блока КМ (рис. 2) представляет собой замкнутый контур, включающий два уровня информационного взаимодействия: первый уровень представлен каналом инкапсуляции знаний и каналом анализа состояния, второй уровень содержит канал репрезентации знаний и канал идентификации состояния обучаемого.

По отношению к процессу формирования знаний субъекты ИОС системы АДО разграничены по уровню доступа, выступая в различных ролях: профицитные единицы (П.Е.): администратор, автор, тьютор; дефицитные единицы (Д.Е.): гость, абитуриент, учащийся. Виртуальный диалог между субъектами осуществляется посредством средств ИОС системы АДО и поэтому обладает существенным недостатком – ограниченностью коммуникативного взаимодействия (в широком смысле), который необходимо технологически устранять.

Обучение рассматривается как процесс управляемого переноса предметных знаний в сознание обучаемого и включает последовательность этапов обработки информации (визуальная репрезентация, восприятие, понимание, формирование навыков, агрегация полученной информации в знания), поэтому уровень остаточных знаний обучаемого зависит от качества восприятия и понимания информационных фрагментов.

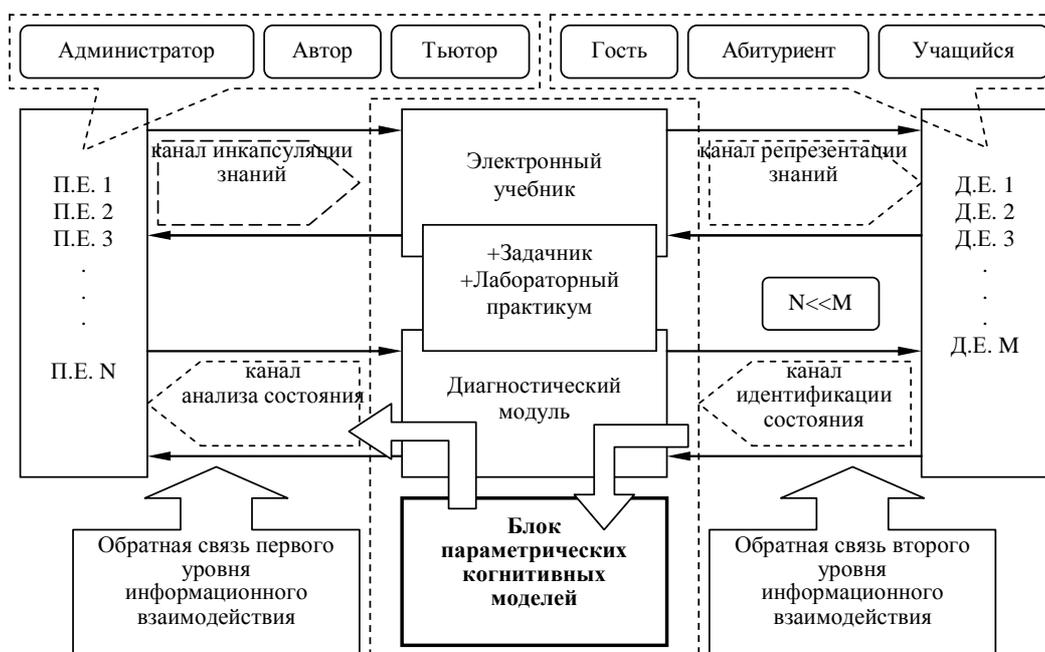


Рис. 2.

Структура системы автоматизированного (дистанционного) обучения с элементами адаптации на основе блока параметрических когнитивных моделей

Диагностический модуль системы АДО (рис. 3) предназначен для начального, промежуточного и итогового контрольного тестирования уровня остаточных знаний обучаемого, а также реализует начальное и квазидинамическое исследование параметров ЛХО для наполнения КМ субъекта обучения.



Рис. 3. Общая схема обучения как управляемого процесса с адаптацией на основе блока параметрических когнитивных моделей: принцип функционирования диагностического модуля

Управление отображением информационных фрагментов (электронная книга, раздел, глава, параграф) различными способами реализует лингвистический процессор электронного учебника (рис. 4) на основе структурной (семантической) метамодели предмета и блока параметрических КМ. Структурная (семантическая) метамодель предмета необходима и достаточна для наполнения предметным содержанием по циклу дисциплин.

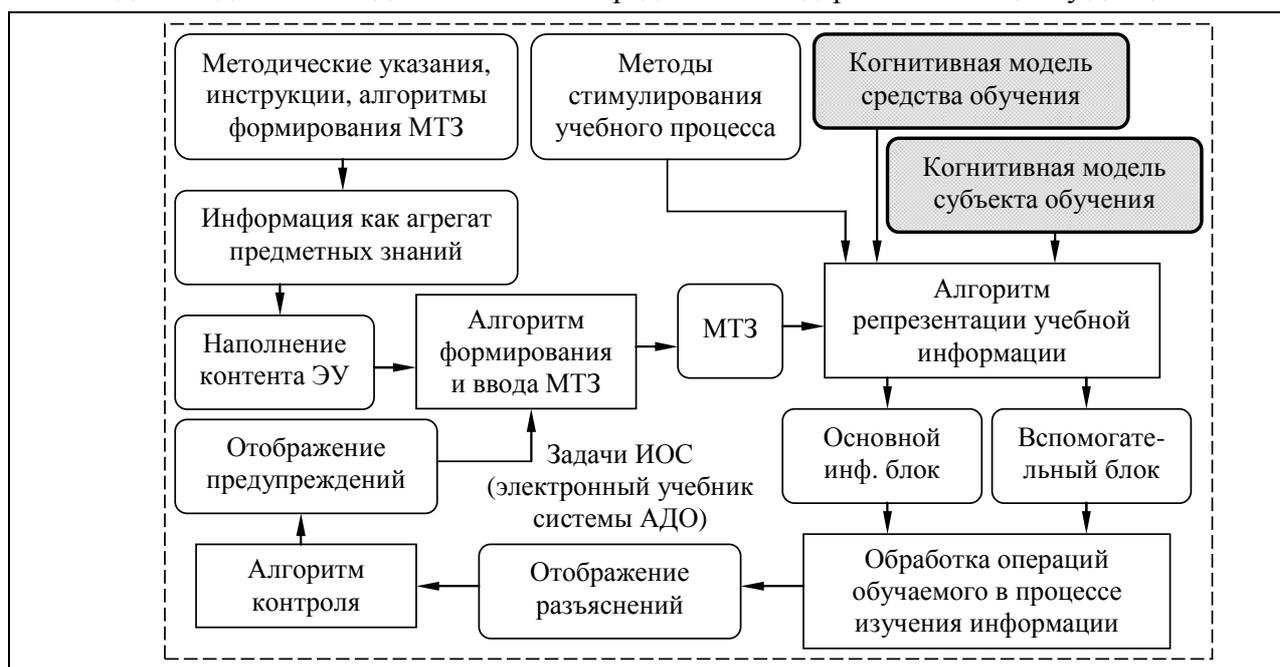


Рис. 4. Общая схема обучения как управляемого процесса с адаптацией на основе блока параметрических когнитивных моделей: принцип функционирования электронного учебника

В электронном учебнике материал по каждой дисциплине структурируется на главы, разделы, подразделы, каждому информационному фрагменту ассоциируется блок вопросов, предназначенный для использования в диагностическом модуле системы АДО, что позволяет эффективно организовать текущий, промежуточный и итоговый контроль уровня остаточных знаний субъекта по ряду предметов изучения (дисциплин) с использованием ряда различных методик оценки.

Для автоматизации исследования уровня остаточных знаний разработан программный инструментарий, позволяющий проводить оценку на основе двух шкал: «грубая» – расчет суммы правильных ответов на вопросы); расширенная – на основе суммы набранных баллов по каждому варианту ответа на вопрос.

Обработка апостериорных данных тестирования осуществляется с использованием специальной методики (не представлена в статье) позволяющей оценить уровень знаний испытуемого и провести анализ качества теста на основе ряда коэффициентов: сложность задания; суммарный результат выполнения заданий i -м учащимся; суммарный результат выполнения j -го задания всеми учащимися; средний уровень тестирования по результатам выполнения всех заданий; средний уровень выполнения j -го задания всеми учащимися; дисперсия суммарных баллов тестируемых; стандартное отклонение суммарных баллов тестируемых; дисперсия результатов тестирования по j -му заданию; стандартное отклонение результатов тестирования по j -му заданию; оценка связи каждого j -го задания с суммой баллов по всему тесту; среднее арифметическое экспертных оценок; стандартное отклонение экспертных оценок; коэффициент корреляции результатов тестирования и независимых экспертных оценок (валидность теста); надежность теста как показатель точности и устойчивости результатов тестирования.

Для комплексного решения проблемы создания и последующего анализа ИОС системы АДО с элементами адаптации на основе блока параметрических КМ предлагается технология когнитивного моделирования (ТКМ) и методика использования ТКМ для формирования структуры КМ.

ТКМ выступает универсальной по отношению к объекту исследования и является итеративным циклом, включающим последовательность этапов, позволяющих не только получить первичные представления в рамках спектра научных аспектов рассмотрения, но и осуществить структурный анализ. ТКМ включает следующие этапы: идентификация (получение первичных представлений об исследуемой ситуации), концептуализация ситуации (концептуальная схема исследуемой ситуации в предметной области), структурирование ситуации (структурированные знания о ситуации в предметной области), формализация (построение первого и второго уровня структуры КМ), структурный анализ (верификация первого уровня структуры полученной КМ), параметрический анализ (верификация второго уровня структуры КМ, ее параметров), реализация (размещение полученной КМ в основе среды исследования), моделирование (моделирование основанное на целостном подходе), анализ (статистическая обработка

полученных с помощью КМ данных), предметная интерпретация (интерпретация полученных зависимостей, знаний), синтез (накопление новых знаний о динамике развития ситуации в предметной области). Для сложных объектов исследования ТКМ предусматривает привлечение консультантов: эксперт – предметный специалист (для ИОС методист и пр.); когнитолог – специалист в области инженерии знаний; системный аналитик – специалист в области системного анализа и моделирования; программист – квалифицированный специалист, владеющий методами и подходами к реализации современных программных средств посредством высокотехнологичных сред программирования.

Методика использования технологии и алгоритм формирования КМ (не представлены в статье) разработаны для формализации последовательности использования ТКМ с целью формирования КМ для задач анализа ИОС АДО с элементами адаптации на основе блока параметрических КМ.

Для исследования структурно сложных объектов рекомендуется использовать представление КМ в виде ориентированного графа, вершины которого образуют ряд множеств (рис. 5, вверху), что является удобным для последующего анализа.

Для структурно простых объектов исследования рекомендуется использовать схематическое представление КМ (рис. 5, внизу).

КМ – параметризованный репертуар, эшелонированный на совокупность портретов и стратифицированный на ряд множеств (представление КМ на рис. 5, вверху). Для ИОС АДО КМ отражает наиболее важные аспекты и параметры информационного взаимодействия субъектов и средств обучения, обеспечивающие согласованность генерации образовательных воздействий с личностными характеристиками (особенностями) субъектов обучения, а также позволяющим выявить причины затруднений в процессе формирования знаний.

Контур адаптации в ИОС АДО на основе блока КМ технологически реализуем при возможности генерации обучающих воздействий на основе КМ образовательного средства согласованно с ЛХО содержащимися в КМ субъекта.

Соответственно КМ дифференцируется на КМ субъекта (параметры, характеризующие ЛХО) и КМ средства обучения (параметры, характеризующие потенциально возможные типы и виды генерируемых информационно-образовательных воздействий).

КМ субъекта обучения (рис. 6) концентрирует параметры физиологического, психологического и лингвистического портретов, характеризующие ЛХО.

КМ средства обучения (рис. 7) содержит параметры, характеризующие потенциально возможные типы, виды и особенности генерируемых информационно-образовательных воздействий.

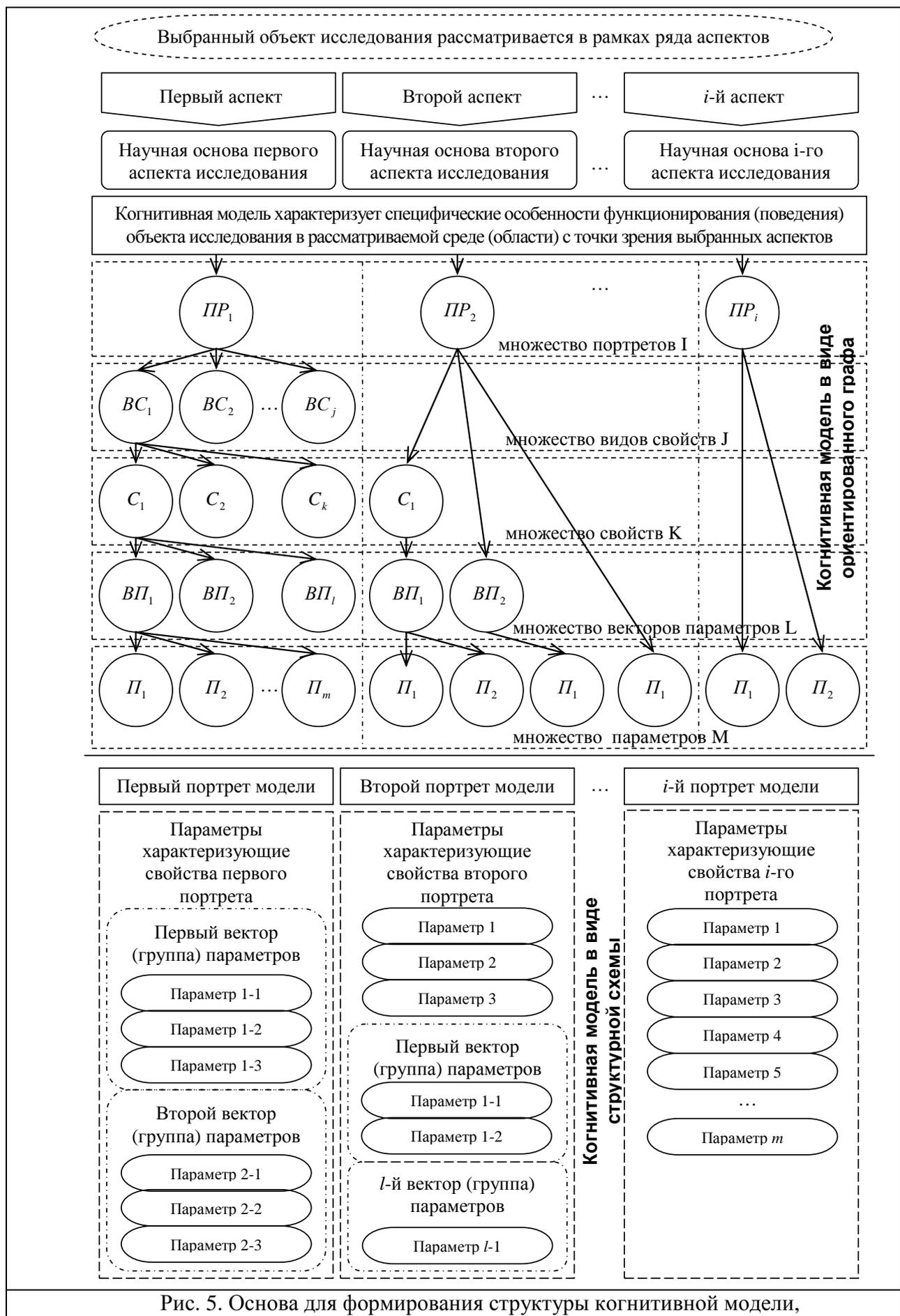


Рис. 5. Основа для формирования структуры когнитивной модели,

представленная в виде графа (вверху) и структурной схемы (внизу)

Испытуемый (модель) рассматривается в спектре трех научных аспектов



Рис. 6. Когнитивная модель субъекта обучения

Средство обучения (модель) рассматривается в спектре трех научных аспектов

Физиологический аспект

Лингвистический аспект

Психологический аспект

Частная физиология
анализаторов

Прикладная
лингвистика

Когнитивная
психология

Когнитивная модель средства обучения характеризует особенности генерации информационно-образовательных воздействий с учетом индивидуальных особенностей субъекта обучения

Физиологический портрет

Лингвистический портрет

Психологический портрет

Параметры визуальной
репрезентации

Параметры фона

тип узора

цвет фона

комбинация цветов

Параметры шрифта

гарнитура шрифта

размер кегля символа

цвет символа

Цветовые схемы

при ахромазии

при протанопии

при дейтеранопии

при тританопии

Параметры звуковой
репрезентации

Особенности воспроизведе-
дения звукового потока

громкость

тембр

тип потока

звуковая схема

Языковые аспекты
коммуникации

уровень изложения
материала

набор ключевых слов
и определений

набор элементов в
основе интерфейса
взаимодействия

Дополнительные
возможности

коррекция последоват.

навигация по курсу

добавление модулей

выбор вида информ.

выбор стиля предст.

выбор скорости предс.

творческие задания

дополнит. модули

дополнит. литература

Скорость репрезентации

быстрая

медленная

Способ
репрезентации

Вид информации

текстовая

табличная

схематическая плоск.

схематическая объемная

звуковая (основная)

звуковая (сопровожд.)

комбинированная

специальная схема

Стиль представления

целостное представление/
детализированное пред.

автоматическое
переключение/ручное

постоянный тип
информации/переменный

глубокая конкретизация /
абстрактное изложение

простота изложения/
сложность изложения

широкий набор
терминов/ узкий набор

Рис. 7. Когнитивная модель средства обучения

Для автоматизации задач исследования векторов параметров в составе портретов КМ субъекта использовался программный инструментарий, разработанный на базе архитектуры экспертной системы и содержащий в основе базы знаний совокупность прикладных методик на научной основе частной физиологии сенсорных систем, когнитивной психологии и лингвистики.

Оценка эффективности внедрения результатов исследования производилась с использованием общепринятых показателей эффективности:

$$K = \{k_1; k_2; k_3\} = \left\{ Y_1 - Y_2; \frac{Y_1}{Y_2}; \frac{Y_1 - Y_2}{Y_2} 100\% \right\}$$

Коэффициенты соответственно обозначают абсолютный, сравнительный и относительный показатели эффективности, а результаты статистической обработки апостериорных данных серии экспериментов обобщены в таблице 1.

Таблица 1. Результаты статистической обработки данных эксперимента

| Наименование показателей | Номер экспериментальной группы испытуемых | | | | | | | |
|--|---|--------|--------|--------|--------|--------|--------|--------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Эксперимент №1 (без использования технологии) | | | | | | | | |
| Средний балл Y1 | 3,850 | 3,414 | 3,224 | 3,678 | 4,036 | 3,643 | 3,790 | 3,645 |
| СКО ср. балла | 0,867 | 0,178 | 1,958 | 0,879 | 0,577 | 0,783 | 1,679 | 1,047 |
| Эксперимент №2 (с использованием технологии, личностная адаптация) | | | | | | | | |
| Средний балл Y2 | 4,041 | 3,674 | 3,357 | 3,786 | 4,157 | 3,853 | 3,821 | 3,743 |
| СКО ср.балла | 0,723 | 0,127 | 1,743 | 0,743 | 0,446 | 0,654 | 1,538 | 0,986 |
| Итоги исследования | | | | | | | | |
| K1 | 0,191 | 0,026 | 0,133 | 0,108 | 0,121 | 0,021 | 0,031 | 0,098 |
| K2 | 1,049 | 1,076 | 1,041 | 1,029 | 1,029 | 1,057 | 1,008 | 1,026 |
| K3 | 0,049 | 0,076 | 0,041 | 0,029 | 0,029 | 0,057 | 0,008 | 0,027 |
| Изменение СКО | -0,144 | -0,051 | -0,215 | -0,136 | -0,131 | -0,129 | -0,141 | -0,061 |

Предложенный подход позволяет реализовать дополнительный контур адаптации на основе блока параметрических когнитивных моделей, а также провести анализ эффективности функционирования информационно-образовательной среды автоматизированного (дистанционного) обучения.

Список литературы

1. Аванесов В.С. Композиция тестовых заданий. – М.: Адепт, 1998.
2. Башмаков А.И. Разработка компьютерных учебников и обучающих систем. – М.: Филин, 2003.
3. Ершов А.П. Концепция использования средств вычислительной техники в сфере образования (информатизация образования). – Новосибирск: Препринт ВЦ СО РАН АН СССР, 1990.
4. Зенкин А.А. Когнитивная компьютерная графика. – М.: Наука, 1991.
5. Каймин В.А. Технология разработки учебных программных средств. – М.: ИНФО, 1987.
6. Кашицин В.П. Системы дистанционного обучения: модели и технологии // Проблемы информатизации.- 1996.
7. Кроль В.М. Психофизиологические аспекты разработки визуального пользовательского интерфейса нового поколения // Пользовательский интерфейс: исследование, проектирование, реализация, 1993, №3.
8. Лобанов Ю.И. Подготовка информации для автоматизированных обучающих систем. – М.: ВШ, 1986.
9. Семенов В.В. Компьютерные технологии в дистанционном обучении. – М.: НИИВО, 1997.
10. Осипов Г.С. Приобретение знаний интеллектуальными системами. – М.: Наука, 1997.

A. N. Vetrov, E. E. Kotova, N. N. Kuzmin

THE ADAPTIVE INFORMATION-EDUCATIONAL ENVIRONMENT OF AUTOMATED (REMOTE) LEARNING ON THE BASE OF PARAMETRICAL COGNITIVE MODELS

The technology of cognitive modeling, methods of its applying and algorithm of forming structure of the cognitive model is suggested for analyzing the efficiency of functioning information environment of adaptive education based on the block of the parametrical cognitive models.

Information-educational environment, automated (remote) education system, cognitive model, module of the parametrical cognitive models, the technology of cognitive modeling, method of constructing of the cognitive model structure