Министерство образования,
науки, молодежи и спорта
Украины
Севастопольский
национальный технический
университет

Российская
академия наук

Институт проблем
информатики

СОВРЕМЕННЫЕ ПРОБЛЕМЫ ПРИКЛАДНОЙ МАТЕМАТИКИ ИНФОРМАТИКИ, АВТОМАТИЗАЦИИ И УПРАВЛЕНИЯ

Материалы
3-го международного
научно-технического семинара
9-13 сентября 2013 г., г. Севастополь

УДК 004

Современные проблемы прикладной математики, информатики, автоматизации, управления // Материалы 3-го международного научно-технического семинара (9-13 сентября 2013 г., г. Севастополь). – Москва. ИПИ РАН, 2013. – 313 с. – ISBN 978-5-91993-025-9

В сборнике представлены материалы докладов 3-го международного научно-технического семинара «Современные проблемы прикладной математики, информатики, автоматизации, управления» (2013 г. Севастопольский национальный технический университет). В докладах изложены перспективные формы и методы решения актуальных задач прикладной математики, информатики, автоматизации и управления.

Материалы докладов публикуются в авторской редакции.

ISBN 978-5-91993-025-9

© Авторы докладов, 2013
© Институт проблем информатики РАН, 2013
© Севастопольский национальный технический университет, 2013

СОДЕРЖАНИЕ

Предисловие ……..... 5

А.Д. Урсул
Образование в «информационном измерении».……..... 10

А.В. Скатков, В.С.Ловягин, С.А.Чорномыз
Информационные технологии критериального обнаружения атак в телекоммуникационных сетях критического применения…… 41

А.А. Брюховецкий, А.В. Скатков, П.О. Березенко
Адаптивная модель обнаружения уязвимостей в критических приложениях на основе решающих деревьев и байесовского классификатора…….. 54

А.Д. Панов
Технологическая сингулярность, теорема Пенроуза об искусственном интеллекте и квантовая природа сознания…….. 63

А.К. Горшенин, С.Я. Шоргин
Разработка информационной технологии интеллектуального анализа больших данных……..... 104

И.Н. Синицын, А.С. Шаламов, В.И. Синицын, Э.Р. Корепанов, В.В. Белоусов, А.А. Кулешов
Методы и средства оценки запасов и уровня готовности систем интегрированной логистической поддержки, основанные на канонических разложениях случайных функций…… 115

И.М. Гуревич
Информационные основы и ограничения нанотехнологий ……..... 127

А.Л. Леонтович, М.П. Евстигнеев, И.Ш. Шенгелая
Некоторые философские, математические и физические аспекты приложений специальной и общей теорий относительности в физике…… 155

А.Б. Филимонов, Н.Б. Филимонов
Ситуационно-целевой подход в задачах автоматизации эргатических систем управления подвижными объектами…… 170

С.Ф. Сергеев
Проблема эффективного взаимодействия человека-оператора с интеллектуальными техническими системами и средами……. 183

И.Ф. Чебурахин
Обработка данных и управление в дискретных системах на основе функционально-логических уравнений……... 198

Ф.Г. Гаращенко, В.Т. Матвієнко
Моделювання та оптимізація динаміки заряджених пучків як задача практичної стійкості……. 216

В.Т. Матвиенко
Оптимизация модальных регуляторов……. 222

Ю.Е. Обжерин
Полумарковские модели систем с резервом времени…… 232

Л.А. Краснодубец, А.А. Кабанов
Конструирование информационной структуры и синтез параметров обратных связей систем траекторного управления мобильными роботами…..... 246

М.П. Фархадов, В.А. Вертлиб, Н.В. Петухова, З.П. Мясоедова
Вопросы создания распределенных систем массовогообслуживания населения с современными речевымиинтерфейсными технологиями……. 254

В.А. Смирнов, М.П. Фархадов
Система анализа неструктурированной речевой информации в сети Интернет……... 276

М.В. Газе, Н.И. Лычагин, А.А. Малаховский
Логическая организация мультисервисных объектовых ведомственных сетей связи…….. 283

И.М. Гуревич, Л.К. Левит-Гуревич
Использование моделей сетевых систем для описания и оценок гидрографических сетей……..... 296

СОВРЕМЕННЫЕ ПРОБЛЕМЫ ПРИКЛАДНОЙ МАТЕМАТИКИ, ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ

Прикладная математика, информатика, автоматизация – важнейшие области современной науки, охватывающие практически все направления человеческой деятельности – прежде всего, исследования, а также проектирование, производство, эксплуатацию и др. Поэтому очередное (третье) издание настоящего сборника представляется весьма важным. Он даст возможность читателям ознакомиться с состоянием, уровнем и перспективами развития данных направлений научных исследований в России и Украине.

Можно выделить четыре группы докладов семинара.

1. Доклады, в которых формулируются и обсуждаются теоретические основы рассматриваемых областей современной науки. Это прежде всего работа А.Д. Урсула «Образование в “информационном измерении”», в которой отмечается, что появление на информационной основе инновационно-опережающих процессов в образовании приведет к существенной его футуризации, становлению опережающего образования и формированию дальнейших эволюционных процессов цивилизации. В докладах А.В. Скаткова, В.С.Ловягина, С.А.Чорномыза «Информационные технологии критериального обнаружения атак в телекоммуникационных сетях критического применения», А.А. Брюховецкого, А.В. Скаткова, П.О. Березенко «Адаптивная модель обнаружения уязвимостей в критических приложениях на основе решающих деревьев и байесовского классификатора» обсуждаются важные на практике проблемы. В том числе, анализируются акты несанкционированного перехвата каналов обмена информацией и управления, атаки и вредоносные программы. Предложены методы использования информационной технологии критериального обнаружения атак и модели обнаружения атак и вредоносных программ. Интересна работа А.Д. Панова «Технологи-

ческая сингулярность, теорема Пенроуза об искусственном интеллекте и квантовая природа сознания», в которой излагаются фундаментальные идеи формирования искусственного интеллекта, моделирования мозга и необходимость уточнения основ физики, причем особое внимание уделяется разъяснению смысла теоремы Пенроуза. Работа А.К. Горшенина, С.Я. Шоргина «Разработка информационной технологии интеллектуального анализа больших данных» посвящена формированию концепции анализа больших данных общего типа, а именно информационных потоков. Существует большой набор областей, в которых поступление данных вполне может соотноситься с определенными событиями, что автоматически формирует информационный поток, подлежащий анализу. В докладе И.Н. Синицына, А.С. Шаламова, В.И. Синицына, Э.Р. Корепанова, В.В Белоусова, А.А. Кулешова «Методы и средства оценки запасов и уровня готовности систем интегрированной логистической поддержки, основанные на канонических разложениях случайных функций» приводятся результаты разработки и испытаний типовой инструментальной системы «Динамика процессов», предназначенной для автоматического поиска и анализа оптимальных поставок запасных частей в системах послепродажного обслуживания изделий наукоемкой продукции. Это позволяет кардинально реформировать традиционные системы управления путем внедрения методов оптимального планирования процессов расходования ресурсов по критериям экономической целесообразности и эффективности. В работе И.М. Гуревича «Информационные основы и ограничения нанотехнологий» определяются фундаментальные информационные константы и на их основе формулируются ограничения, накладываемые ими на возможности нанотехнологий. В работе А.Л. Леонтовича, М.П. Евстигнеева, И.Ш. Шенгелая «Некоторые философские, математические и физические аспекты приложений специальной и общей теорий относительности в физике» показано значение одинаковости законов физики в любых системах отсчета, из которой следует принципиальная познаваемость мира, основанная на регистрации и анализе физических сигналов,

исходящих от исследуемого объекта.

2. Работы, посвященные вопросам управления и оптимизации сложных технических систем. В работе А.Б. Филимонова и Н.Б. Филимонова «Ситуационно-целевой подход в задачах автоматизации эргатических систем управления подвижными объектами» обсуждается методология автоматизации эргатических систем данного класса, основанная на иерархической декомпозиции функциональных задач, центральное место в которой занимает концепция ситуационно-целевого управления. С.Ф. Сергеев в работе «Проблема эффективного взаимодействия человека-оператора с интеллектуальными техническими системами и средами» рассматривает возможные варианты симбиотического объединения интеллекта человека-оператора с интеллектом технических систем и сред. В докладе И.Ф. Чебурахина «Обработка данных и управление в дискретных системах на основе функционально-логических уравнений» обсуждаются методы реализации логических алгоритмов (булевых функций) на основе структурно-функциональной параллельной (логической) декомпозиции. В работе Ф.Г. Гаращенко и В.Т. Матвиенко «Моделирование и оптимизация динамики заряженных пучков как задача практической устойчивости» исследуются тесно связанные задачи устойчивости и оптимального управления потоками частиц. Оптимизация систем в структурно-параметри-ческом классе позволило разработать ряд алгоритмов получения квазиоптимальных решений, которые зарекомендовали свою работоспособность в реальных установках. В работе В.Т. Матвиенко «Оптимизация модальных регуляторов» рассматривается задача оптимального выбора структуры распределения управляющего сигнала в обратной связи линейной системы. В докладе Ю.Е. Обжерина «Полумарковские модели систем с резервом времени» предлагаются методы обеспечения нормального функционирования систем в условиях воздействий внешних возмущений путем назначения и использования резервного

(избыточного) времени. В работе Л. А. Краснодубца, А. А. Кабанова «Конструирование информационной структуры и синтез параметров обратных связей систем траекторного управления мобильными роботами» рассматривается проектирование законов траекторного управления МР на основе методов оптимизации и обратных задач механики.

3. Работы, посвященные системам массового обслуживания. В работе М.П. Фархадова, В.А. Вертлиба, Н.В. Петуховой, З.П. Мясоедовой «Вопросы создания распределенных систем массового обслуживания населения с современными речевыми интерфейсными технологиями» предложены оптимальные решения для объединения разнородных систем обслуживания населения на базе использования принципов создания сетевых структур дистрибутивно-инвенторного типа и унифицированного внутрисетевого взаимодействия. В.А. Смирнов и М.П. Фархадов в работе «Система анализа неструктурированной речевой информации в сети Интернет» рассматривают принципиальную схему развертывания системы анализа неструктурированной речевой информации в сети Интернет на примере системы поиска ключевых слов в аудиозаписях.

4. Работы посвященные сетевым системам. В работе М.В. Газе, Н.И. Лычагина, А.А. Малаховского «Логическая организация мультисервисных объектовых ведомственных сетей связи» предлагается осуществлять комплексное оснащение различных уровней АСУ отечественным оборудованием, ориентированным на строгое соблюдение перспективных международных стандартов и рекомендаций, имеющих четкую документированную системную идеологию, распространяющуюся на все уровни иерархии АСУ, и обладающих возможностями адаптации к требованиям различных ведомств по обеспечению информационной безопасности, включая работу в различных операционных средах. И.М. Гуревич и Л. К. Левит-Гуревич в работе «Использование моделей сетевых систем для описания и оценок гидрографических сетей» обосновывают возможность

использования многоуровневых статических и динамических моделей сетей для оценок характеристик разветвленных водных систем.

Настоящий сборник будет полезен студентам, аспирантам, преподавателям ВУЗов и ученым России, Украины и других стран, получающим образование и работающим в различных областях современной науки и техники.

Ректор СевНТУ профессор Е.В. Пашков Директор ИПИ РАН академик И.А. Соколов

УДК 681.32
А.Д. Урсул, д-р филос. наук, проф.,
академик Академии наук Молдовы, засл. деятель науки РФ
Центр глобальных исследований МГУ им.М.В. Ломоносова.
г. Москва, Россия
ursul-ad@mail.ru

ОБРАЗОВАНИЕ В «ИНФОРМАЦИОННОМ ИЗМЕРЕНИИ»

Введение

В дальнейшем изложении я буду исходить из концепции всеобщности информации, которую развиваю уже в течение почти полувека [1]. Характерно, что к всеобщему характеру информации тяготеют в основном естествоиспытатели, особенно представители наук о неживой природе, а к социоцентрическому видению информационных процессов - специалисты в области социально-гуманитарного знания. Эти представители двух культурных традиций пока никак не могут договориться между собой не только по вопросу о природе информации, но и по ряду других вопросов, на что в свое время обратил внимание физик и писатель Ч.П. Сноу [2].

В самом общем виде предполагается, что информация, также как и энергия, существует во всех сферах и фрагментах мироздания, является характеристикой всех материальных систем и форм существования материи в нашем мироздании. При таком подходе при рассмотрении взаимодействия материальных объектов (систем) между ними происходит обмен не только веществом и энергией, но и информацией. Если акцентировать внимание только на информационном аспекте этого взаимодействия, то его можно рассматривать как «расширенный» отражательно-коммуникативный процесс, выходящий за пределы социума. Этот процесс в общем виде характеризует не только

общение между людьми, либо между любыми живыми существами, но также и обмен информацией между ними и объектами неживой природы. Такой обмен как взаимодействие и взаимотражение материальных систем реализуется и в сфере неживой природы, хотя в ряде случаев взаимодействие «вырождается» в одностороннее воздействие одного объекта (скажем, темной энергии на все остальные фрагменты Вселенной).

Информацию, в широком понимании, кроме отражения (коммуникации), связывают с разнообразием и различием или неоднородностью. С моей точки зрения, информация - это объективная и системная характеристика не только отражения, но и разнообразия. В этом существенное отличие информации от энергии: при использовании энергетического подхода абстрагируются от наличия разнообразия и неоднородностей в мире. Поэтому все объекты (системы) с точки зрения энергетического подхода различаются только количеством заключенной в них энергии (массы). Я исхожу из вывода, сделанного мною почти полвека тому назад, что информация в самом общем виде выступает как некое системное единство отражения и разнообразия, или как разнообразие отражения. Причем пока этот вывод не только выдержал проверку временем, но и был существенно обоснован в основном благодаря использованию этой концепции в науках о неживой природе и междисциплинарных исследованиях (особенно - в синергетике).

Благодаря осознанию роли информации и информационных процессов, человечество находится в состоянии глобальной информационной революции, связанной с переходом от печатных форм хранения, передачи и других форм хранения и движения информации к электронной форме. В результате этой научно-технологической революции появляется глобальное информационное общество, которое следует характеризовать не только выдвижением на приоритетное место информации по отношению к веще-

ству и энергии, но и последующей преимущественно информационной ориентацией и стратегией всей его деятельности.

Образование, использующее новые информационные технологии, должно стать ядром информационного общества и одним из приоритетных механизмов его дальнейшего развития, особенно при переходе к глобальному устойчивому развитию и последующему ноосферогенезу. А это вместе с тем означает усиление внимания ко всем информационным аспектам образовательного процесса. Расширение и фундаментализация категориального статуса информации позволяет по-новому рассматривать ряд проблем, в числе которых такой феномен как образование.

В статье пойдет речь в основном о некотором новом видении феномена образования при использовании информационного подхода в эволюционном ракурсе. Будет предложено понимание образовательного процесса, которое выходит за рамки принятого в педагогических науках. Такое видение образования в более широком аспекте будет содействовать выработке адекватного системного образа этого феномена в жизни человека, эволюции общества и природы.

Информационный подход к образованию в большинстве работ сводится к исследованию применения новых информационных технологий в этом процессе и есть также работы в этом плане, относящиеся к управлению образованием. Однако очень мало работ, представляющих образование как информационный процесс, тем более в широком эволюционном ракурсе. Между тем образование по своей природе имеет информационный характер и такое представление и моделирование может раскрыть его сущность и смысл как в ходе социальной эволюции, так и в более широком понимании этого феномена.

Информационная модель образования

Далее будут рассмотрены некоторые соображения в

плане формирования концептуальной модели образования как одного из важнейших информационных процессов в обществе. Образование как социально-информационный процесс, включающий в себя триединство (обучение, воспитание, развитие) берет свое начало в предыдущей – биологической ступени, которой присуще такое фундаментальное свойство как обучение, которое характерно для всех животных (причем в своих зачаточных формах свойственно даже одноклеточным). Обучение рассматривается и в кибернетике, где это понятие трактуется как информационный процесс, при котором адаптивная кибернетическая система развивает (совершенствует) свои характеристики под воздействием внешнего разнообразия.

Обсуждаемые далее подходы так или иначе связаны с определенным видением информационной модели образования. Образование может моделироваться информационными методами, поскольку в определенном аспекте действительно представляет собой информационный процесс накопления, передачи, переработки и других форм движения особого рода информации – образовательной информации. В этом аспекте образовательный процесс может моделироваться передачей информации от учителя к ученику (и наоборот), от одного поколения людей к другому, а также различными формами сохранения и преобразования информации, ведущими к появлению определенных социальных результатов.

На уровне социальной ступени эволюции обучение представляется как целенаправленный организованный процесс передачи и усвоения знаний, умений, навыков и способов включения человека в социальную деятельность. Это двустороннее взаимодействие, в котором участвуют как обучающий, так и обучаемый в их совместной информационно-ком-муникативной деятельности.

Особенностью любого процесса развития, включая образование, является появление нового, которое может возникать наряду с тем, что уже существовало, т.е. совме-

стно со старым, традиционным, так и при условии изменения или даже исчезновения старого. Возникновение нового в эволюционных процессах уже заметно на уровне исследования феномена развития, еще без проникновения в его сущность и оно нашло свои интерпретации, в частности, при использовании информационного подхода, когда констатируется, что соединение элементов в некоторую целостность дает приращение новой информации в уже имеющейся системе и создает эффект новой целостности [3].

Причем к образованию, как и к любому процессу развития, применим информационный критерий развития, связывающий увеличение степени (уровня) развития материальных систем с ростом их информационного содержания. Информационный критерий представляет способ оценки изменения информационного содержания (и связанной с ним негэнтропии) материальных систем в ходе эволюционной самоорганизации либо самодезорганизации любой системы, включая образование. Причем на прогрессивной линии эволюции происходит непрерывное накопление информации в системах, и тем самым этот критерий выступает в качестве вектора и уровня развития материальных систем в широком смысле.

Этот критерий в образовании выражает достаточно очевидную векторно-генетичес-кую связь роста информационного содержания в обучаемом (обучаемых) как одной из подсистем, а, скорее всего, в обеих образовательных подсистемах, но в разных направлениях и измерениях. Этот критерий имеет также синергетический смысл, так как этот рост информационного содержания у обучаемого происходит в основном за счет направляющей негэнтропийной активности учителя. Или, в более широком смысле, за счет окружающей информационной социальной и природной среды, из которой в разных ситуациях черпаются информационные ресурсы и которая представляет информационно-экологические условия существования обучаемого как открытой самоорганизующейся системы.

При передаче и движении информации в образовательном процессе она обладает всеми свойствами, которые уже выявлены в информатике и других науках об информации, в частности, такими как синтаксические, семантические и прагматические характеристики. Эти семиотические свойства «образовательной информации», которая циркулирует в процессах и системах обучения, развития и воспитания, выражают формально-количественные параметры, значение и ценность этого вида информации. Я не буду здесь рассматривать отличие информации от знания и отличие образовательного знания от научного знания, поскольку эти проблемы уже неоднократно рассматривались в литературе. Имеет смысл обратить внимание как относятся упомянутые общие свойства информации как образовательного знания к предмету образовательной деятельности, которые обычно представлялись как знания, умения и навыки (ЗУН). Вместе с тем уже предлагается иная триада – знание, понимание, умение – ЗПУ [4, 5], что означает акцент на понимании как семантической характеристике образовательной информации.

В информационно-образовательном процессе в той или иной форме проявляются все основные семиотические (синтаксические, семантические и прагматические) характеристики образовательного знания, хотя ранее они получили иные названия. Само по себе это знание уже содержит все основные характеристики (свойства) информации, и в зависимости от конкретного образовательного процесса они проявляются в том или ином виде и отношении. Если за триаду принимается традиционное ЗУН и акцент делается на определенных прагматических свойствах информации (позволяющих их использовать в соответствующих видах практик), то такая ситуация характерна для утилитарно-профессионального понимания образования, поскольку знания в основной предметной триаде дополняются умениями и навыками (компетенциями). И такое видение предметности образовательной деятельности в опреде-

ленных ее видах и формах вполне приемлемо для некоторых направлений специального среднего и высшего профессионального образования, что специально подчеркивается в ныне пропагандируемом компетентностном подходе, в центре которого - умения.

Впрочем, речь далее пойдет не только о компетентностном аспекте процесса социализации, но и более широком информационном процессе взаимодействия человека с окружающей его социальной и природной средой. И этот процесс взаимодействия в самом общем случае также можно описать (моделировать) информационно-разнообразностным (синтаксическим) методом, предложенным У.Р. Эшби [6], сформулированным им в форме закона необходимого разнообразия.

Феномен футуризации образования: информационный ракурс

Образование как специфический социальный институт формирования и социализации личности с помощью обучения и воспитания возникает на определенном и довольно позднем этапе человеческой истории (в так называемое осевое время). В Древней Греции на рубеже V в. до н.э. появляются первые школы, где в качестве учителей выступают философы, ведущие со своими собеседниками-учениками диалоги и читающие им лекции. Так, в IV в. до н.э. в первой перипатетической школе (школе Аристотеля) образование афинской молодежи сочеталось с систематической разработкой научных (философских) знаний. Философские учения создавались во взаимодействии учителей с их учениками, это был единый синкретический научно-образовательной процесс (если пользоваться современной терминологией).

Здесь важно обратить внимание на то, что одновременно с выделением и расширением сферы специально организованного образования постепенно происходило его отставание от жизни общества, от того, что именуется про-

исходящим [7]. Специализированная образовательная деятельность, следуя своим внутренним механизмам и логике саморазвития, сознательно, благодаря усилиям многих поколений педагогов, установила правила, согласно которым образование принципиально было обречено на информационное отставание от бытия и от передовой науки. Наличию «информационного лага» между наукой и образованием не придавалось особого (и тем, более важного) значения, поскольку это считалось (и до сих пор считается) само собой разумеющимся.

Подобный акцент на прошлом стал «аксиомой» педагогической деятельности: учебники, учебно-методические пособия, государственные образовательные стандарты, программы и т.д., как выше подчеркивалось, ориентируются на «устоявшиеся» знания и другие аналогичные информационные продукты. Если в современной науке (а я имею в виду прежде всего общественные науки) свыше 95% знаний содержат информацию о прошлом, то «образовательное знание» повышает этот процент почти до ста. Схематизируя, можно сказать, что роль педагога в подавляющем большинстве случаев сводилась к «передаточному звену» уже устаревшей информации, а учащихся – к роли получателей этой информации, которую они должны были запоминать.

Подобная схема педагогического процесса, конечно, весьма груба и упрощена, но важно показать, что именно подобная схема делала акцент не на главные составляющие сознания человека – стремление человека к новизне, которые имело смысл развивать в процессе формирования личности. Современное образование, делая акцент на прошлом и его запоминании, оказывается фактически отторгнутым от созидания будущего, и такое противоречие необходимо разрешить на пути адаптации инновационных процессов в образовании к будущему. Это означает, что целесообразно в ходе дальнейшего развития упомянутых инновационных процессов устранить в существенной степени темпораль-

ную асимметрию прошлого и будущего по отношению к настоящему, а также начать формировать сознание человека в ином – оптимальном направлении.

Более того - до сих пор подавляющая часть отечественных философов полагает (и учит этому студентов), что сознание в принципе отстает от бытия (и в принципе должно отставать) и это до недавнего времени возводилось даже в ранг одной из закономерностей социальной философии. Но, если именно образование играет решающую роль в формировании сознания человека, а оно отстает от бытия, то образование, следуя этой «логике», должно ориентироваться на прошлое, в нем будут «циркулировать» только устаревшие знания, умения и навыки. Да и понимание содержания образовательного знания нацелено в основном на прошлое и лишь отчасти - на происходящее, во всяком случае, если при этом преследуется цель овладения уже имеющимися истинами.

Ведь в будущем пока нет ни истин, ни фактов, однако есть какая-то информация, способствующая их появлению и освоению наукой и образованием. Как видим, если в науке, а затем и в образовании делается акцент на генерацию и освоение истин, то акцент на прошлом оказывается вполне естественным и логичным. Информационное отставание (лаг) от жизни и науки тем самым изначально был заложен в образовательный процесс. Причем фактическое подтверждение теперь уже сформулированного теоретического положения уже давно хорошо известно [8, 9]. Установлено, что все мировое образование оказалось весьма консервативной социальной системой, которая не удовлетворяет потребностей современной жизни и в основном моделирует в существенно деформированном виде прошлое нашей цивилизации и ее науки [10].

Существование «отстающего образования» в модели неустойчивого развития (т.е. в прошлом и настоящее время), формирует все же либо в основном «отстающее» от жизни личностное (а тем более, общественное) сознание,

либо ее сиюминутное отображение. Понятно, что ориентатация образования на принцип «здесь и сейчас», тесно связанная с происходящим, с «обучением жизнью» представляется неадекватной для образования будущего, оказывается несовместимой с выживанием человечества и противоречит переходу к устойчивому развитию (УР) как принятой мировым сообществом стратегии выживания человечества в условиях обострения глобальных проблем и опасностей.

Формирование сознания человека, причем такого сознания, которое может содействовать переходу к УР, оказалось тем самым под угрозой. Отстающее от бытия сознание - это удел отдельного человека и всего мирового сообщества в модели неустойчивого развития. А другого пути перехода к УР как «через опережающее сознание» просто не существует. Необходимо сформировать сознание, опережающее бытие у «критического количества» населения планеты, без чего невозможен переход к УР. Это возможно и необходимо сделать лишь с помощью образования, которое повернет сознание индивида и подавляющего большинства населения планеты в сторону будущего, прежде всего в его нормативно-устойчивой и глобальной форме.

И здесь возможны два пути его формирования, которые выявляются на пути осмысления общего феномена глобализации и глобализации информационных процессов, а в перспективе - становления общепланетарного общества знания и цивилизации образования. Глобализация, устанавливая различные связи в социуме (а также общества и природы), близкие к консенсусно-коэволюционным, развертывается в нескольких общих формах. Одна из них, на которую чаще обращают внимание – это формирование различного рода связей между фрагментами общества, это своего рода, как иногда говорят, «интеграция различий» (межнациональные, межгосударственные и тому подобные взаимодействия), что ведет к обретению целостности ци-

вилизации.

Вместе с тем существует и другая форма развертывания этого социально-глобального процесса – движение каких-то универсалий, общечеловеческих целей и ценностей, норм, идеалов, общих форм и способов развития (например, рыночных отношений, демократии, прав человека и т.п.) – такая форма может быть названа транснационализацией, или универсальной глобализацией. Обе формы – транснациональная и интегративная глобализации ведет при складывающихся благоприятных обстоятельствах к общепланетарному единству цивилизации.

Такого рода формы и способы глобализации (интеграционная и универсально-транс-граничная) характерны и для науки, особенно для междисциплинарных, трансдисциплинарных и интегративно-общенаучных исследований, и для формирования глобального опережающего сознания. Это последнее будет формироваться и в ракурсе интеграции различных знаний и мнений, форм и способов познания мира, что условно можно обозначить как процесс глобализации науки и образования, что будет вести к глобализации как интеграции различных форм сознания. Вместе с тем все большую интенсивность набирает форма уже не глобализации, а становления глобального сознания через развитие глобальных исследований и появления особой и очень важной формы знания – глобального знания. Хотелось бы обратить внимание на то, что появление этой глобальной формы знания оказалось не просто необходимым, а даже более важным, чем процесс глобализации сознания. Это связано с тем, что глобальное знание намного быстрее формирует глобальное сознание, чем процесс его глобализации, на пути которого стоят очень большие трудности антиинтегративного и негативного характера. Вполне понятно, что все эти процессы имеют информационную природу, на что я уже обратил внимание [11].

Глобализация знания и формирование глобального знания в науке и образовании – это важные информацион-

ные процессы, ведущие к ускоренному накоплению информации и новым – глобальным алгоритмам её упорядочивания и систематизации. Причем генерация глобального знания в науке и использование его в образовании будет происходить гораздо быстрее, чем глобализация науки и образования, поскольку это оказывается наиболее эффективным путем формирования глобального сознания и общепланетарного информационного общества, в котором так нуждается человечество на пути преодоления глобальных кризисов и перехода к УР как стратегии выживания всего человечества. Это объясняет, почему глобальные исследования и глобалистика, в частности, развиваются гораздо быстрее многих других научных исследований, а глобальное образование уже за несколько десятилетий находится в неплохом стартовом положении, захватывая территории мирового образовательного пространства. Глобализация науки и глобальное образование существенно усиливают современные информационные процессы и способствуют ускоренному накоплению и преобразованию наиболее важной для существования и устойчивого развития человечества информации.

Другого пути перехода к УР, кроме формирования глобального опережающего сознания у той части населения планеты, которое принимает глобально-международные и государственные решения, не существует. Сознание в этом случае постепенно и в какой-то своей части станет опережать бытие, и будет формироваться совершенно новая форма принятия эффективных и опережающих решений на этих уровнях, которую я связываю со становлением такой суперглобальной информационной системы как интегральный ноосферный интеллект, в который будет входить научно-образовательный процесс. Появление последнего становится возможным благодаря процессу информатизации, делающему научную мысль планетарно-эволюционным феноменом.

Ноосферный интеллект соединит в одну целостную

социотехнологическую систему культуру как коллективный интеллект цивилизации, естественный интеллект отдельных личностей и искусственный интеллект новых информационных технологий. Эта целостная и самая грандиозная интеллектуальная система обретет способность кардинально увеличить информационные возможности человечества за счет наиболее рационального использования накопленного социокультурного разнообразия.

На пути к ноосферно-опережающему интеллекту

Ноосферный интеллект станет принципиально новой формой «общественного сознания», объединяющей интеллектуальные ресурсы человечества и средства информатики, способной к опережающему отражению социального бытия и использующей опережающие информационные механизмы для реализации желаемых вариантов социоприродного развития. Очевидно, что одно из важных качеств этого интеллектуального фактора будет заключаться в их темпоральных характеристиках. Общественное сознание в социосфере (в модели неустойчивого развития) отстает от социального и социоприродного бытия, позволяя ему развиваться стихийно. В ноосфере же целостно-планетарный интеллект получает возможность предвидеть это бытие, прогнозируя и проектируя и оптимизируя желаемое будущее (будущее, котороое мы хотим [12]).

По мере становления общества с УР трансформируемая реальность также начнет участвовать в формировании опережающего глобально-интегрального сознания, поскольку оно будет ориентировано не только на прошлое и происходящее, но уже и на будущее, причем со все большим нормативным «горизонтом прогноза». Поскольку свойство опережения в образовании (да и фактически во всех других формах деятельности) непосредственно связано не только с экологическими, но и с информационными процессами и проблемами, то в перспективе речь идет о становлении, прежде всего информационного общества с

УР как первого этапа и составной частью будущей сферы разума – инфоноосферы.

Инфоноосфера - будущего информационного общества с УР должно функционировать как первый и вначале основной компонент сферы разума и в определенном смысле должна быть опережающей системой. Выход из современного глобального кризиса цивилизации тем самым видится в эволюционном переходе на путь становления информационного устойчивого общества как одной из ступеней сферы разума, а выход из глобального кризиса образования, в соответствии с этой общей стратегией – в трансформации «отстающей», консервативной образовательной системы в опережающую, или ноосферную образовательную систему. Без информатизации формирование такого качественно нового образования невозможно и поэтому наиболее радикальные инновационные процессы, как можно предположить, связаны с переходом к информационному обществу, являющемуся важной чертой глобального перехода на путь УР. Информационному обществу с УР, базирующемуся на знании и информации, должна соответствовать новая опережающая инновационно-образовательная система.

Я опять не случайно акцентирую внимание на информационном аспекте образования: свойство опережения возникает в ней именно как качественно новый механизм ее формирования. И здесь информационные факторы оказываются приоритетной составляющей грядущей устойчивой цивилизации. В этом смысле свойство «опережения» оказывается присущим не только образованию. В процессе эволюции жизни появилось опережающее отражение, которое, на мой взгляд, существует и в неживой природе, во всяком случае, в вещественной, наблюдаемой Вселенной предположил, что оно формировалось в первые мгновения после Большого Взрыва, когда началось инфляционное расширение Вселенной из начальной космологической сингулярности [13]. Полагаю, что существуют определен-

ные онтологические (космологические) основания процесса футуризации, который проявляется на всех уровнях и ступенях эволюции материи. Онтологические основания процесса футуризации как все большего «расширения» в будущее проявляются и будут все больше проявляться в различных сферах деятельности. Это характерно и для будущего образовательного процесса, который должен обрести кардинально иное инновационное содержание и за счет процесса футуризации [14].

Именно благодаря появлению этого свойства в информационно-отражательных процессах живого вещества стало возможным выживание и эволюция многих высших видов животных. По-видимому, появление и использование различных видов социального опережающего отражения, к которым можно в определенной степени отнести науку, образование, управление, станет информационной базой выживания человеческого рода (выживание через футуризацию социальной деятельности).

Собственно свойство «опережения» будет характерно для всей грядущей эпохи становления ноосферы. Появление этой последней из социосферы связано не только с тем, что на приоритетное место выдвинется разум, но и с тем, что это будет «опережающий» интеллект не только отдельных индивидов, но и всей ноосферы в целом. Ноосфера будет отличаться от социосферы не только огромным информационным потенциалом, но и тем, что этот последний будет структурирован в форму ноосферного интеллекта и, эффективно функционируя, сможет опережать, предвидеть, прогнозировать и управлять «ноосферной» деятельностью уже не столько homo sapiens, сколько уже - homo noosphericus. Все области такой деятельности станут опережающими – научные исследования (не только фундаментальные) будут опережающими, образование в его ноосферной версии также ускоренно будет футуризироваться, управление станет функционировать в упреждающем режиме, формируя интеллектуально-

информационную базу образовании для УР и щего ноосферогенеза.

Уместно обратить внимание на то, что появление свойства опережения как характеристики информационного содержания ноосферы появляется благодаря огромному накоплению информации и появления какой-то части избыточной информации, не используемой для сиюминутного выживания цивилизации. Только наличие такого «запаса» информации позволяет направить часть его на обслуживание опережающих потребностей системы (в данном случае – ноосферы) и тем самым обеспечивать темпоральную оптимизацию её выживания, дальнейшего существования и развития в нерегрессивном направлении.

Приходится менять курс развития всего человечества и формировать концепцию спасения цивилизации на планете, которой, на мой взгляд, выступает концепция УР. Человечество из современного своего неустойчивого состояния должно будет перейти через УР к такому будущему, которое обеспечит его выживание и на этой «устойчивой основе» неопределенно долгое существование. Цель перехода на путь УР состоит не просто в реализации новой формы социального и социоприродного развития, а прежде всего в том, чтобы обеспечить дальнейшее перманентное существование и сохранение человечества как пока единственного вида социальной ступени развития во Вселенной.

Вполне возможно, что если бы динамика глобального развития не имела негативную ориентацию, если бы человеческому роду не угрожала гибель, то «спасительная» функция образования, как и многих (если не всех) сфер деятельности могла быть и не замечена. И проблема понимания и осознания грозящей цивилизации глобальной опасности оказалась бы вне поля зрения исследователей, занимающихся проблемой образования. Между тем акцентирование внимания на проблеме понимания в образовании оказалось весьма тесно связанной с проблемой выхода

из планетарного кризиса на пути перехода к УР.

Начало XXI в. ознаменовалось интенсивным поиском новых моделей образования, которые могли бы удовлетворять потребности не только нынешних, но и будущих поколений землян и всего человечества, поставившего цель своего выживания и сохранения окружающей природной среды [15]. Одну из новых моделей образования мировое сообщество связывает теперь со стратегией устойчивого развития (УР), к которой мировое сообщество решило переходить в XXI в. Основные контуры этой модели социоприродного УР III тысячелетия и его важнейшей подсистемы – образования должны сформироваться уже в ближайшие десятилетия [16]. Совершенно очевидно, что наше общее устойчивое будущее вначале может быть построено как некая его информационная опережающая модель. И дальнейшие действия по ее воплощению в реальность носят, прежде всего, информационный характер. Поэтому переход к УР зависит от создания инновационно-опережающих моделей в любой сфере деятельности и прежде всего в науке, образовании и управлении.

В образовательном процессе должна передаваться информация не только от прошлых поколений к нынешним, но и совершенно новая информация от пока не существующих будущих поколений к настоящим (в форме виртуальных моделей исследования будущего), а также информация о будущем, исходящая уже от нынешних поколений. В этой предлагаемой инновационно-информационной модели образовательного процесса прошлое, настоящее и будущее будут соединены в одну системно-темпоральную целостность. Вопрос, однако, возникает как получить эту новую информацию из будущего и о будущем?

Сразу хочу заметить, что подобная информационно-образовательная модель противоречит не только традиционной педагогической науке, но и вообще современной науке, которая опирается в своей основе на факты, практи-

ку, истину и другие подобные понятия, которые составляют фундамент научно-исследовательского процесса. Совершенно очевидно, что знания о будущем не могут содержать истинных знаний и фактов, которые проверяются практикой, реализующейся лишь в прошлом и настоящем. Будущее в этом смысле отторгнуто от официальной (фактологической) науки, хотя предсказание и прогнозирование и признается функцией теоретического уровня познания. Но, если внимательно приглядеться к тому, что пишется в этом случае об исследовании будущего, то фактически все сведется к линейному видению темпорального процесса (экстраполяции исторических тенденций в будущее).

Это означает, что в исследовании будущего и его внедрении в образовательный процесс должны будут реализовываться новые принципы, способы и формы осознания и освоения грядущего. Одна из наиболее активно развивающихся форм футуризации – это моделирование того, что может произойти в будущем, т.е. виртуально-опережающее моделирование в разных его формах и особенно с помощью новых информационных технологий. Это тот способ вхождения будущего в инновационно-образовательный процесс, который сопряжен с информатизацией, поскольку только на информационном уровне можно строить и изучать модели будущего как нормативные (типа стратегии УР), так и исследовательские прогнозы и предсказания. Здесь возможны как линейные выходы в будущее из прошлого и настоящего, так и нелинейно-инновационные прорывы вплоть до того, что могут создаваться модели, не имеющие прототипов в прошлом или аналогов в настоящем.

И если обучение прошлому основано на запоминании, то включение в этот процесс инновационно-опережающего моделирования базируется на принципах более творческого процесса, развивающего образования. Опережающее образование, которое в принципе является развивающим образованием, было предложено для реали-

зации образования для УР (ОУР) и тем самым принципа темпоральной целостности, причем оказалось ядром образования в интересах устойчивого развития, наряду с экологическим образованием.

Между тем при переходе к УР, прежде всего, должны произойти кардинальные изменения в соотношении информационно-духовных факторов и механизмов цивилизационного развития. Доминирование социально-информационных процессов по отношению к вещественно-энергетическим – это кардинальная трансформация в сознании и нравственности людей, в общечеловеческих ценностях и видении смысла человеческой эволюции, которая отныне должна быть связана со становлением сферы разума. Но это возможно только в том случае, если появится возможность приоритетно-опережающего развития информационных процессов и сопровождающих их развитие социальных механизмов, структур, направлений деятельности, имеющих информационную природу, т.е. науки, образования, духовной культуры, процессов информатизации и информационного обеспечения. Новая парадигма цивилизационного развития оказывается вместе с тем стратегией становления первой ступени ноосферы – информационного общества с УР как инфоноосферы. Это состояние общества в отдельной стране или человечества в целом, которое наступит в результате полной информатизации и основанной на ней интеллектуально-гуманистической перестройки социальных структур и отношений на основе принципов и целей социоприродного УР. Информационное общество в этом случае может в определенной степени стать и обществом знания и обществом образования, поскольку предполагает ускоренное опережающее развитие науки и образования, других сфер духовной культуры, способных к футуризации своего содержания (что в меньшей степени касается других близких сфер реальной жизни в принципе имеющих консервативный характер как, например, большинство религий).

Информационная природа образования: глобально-эволюционные перспективы

Совершенно очевидно, что ОУР и опережающего образование - это разные типы и формы инновационных процессов в образовании. Опережающее образование включает в себя футуризацию и опережающие механизмы для любых вариантов видения будущего, а не только нормативного УР. И в этом смысле оно может использоваться и вне ОУР, и в любой другой учебной дисциплине, курсе и направлении образования, поскольку это инновационный процесс, акцентирующий внимание на будущем. В этом смысле специального опережающего образования такого же как, например, экологическое и экономическое виды образования быть не может. Опережающее образование – фундаментально-универсальная сторона, аспект образования, хотя и весьма слабо развитый в мировом образовании. Однако, если есть в образовании знание о прошлом, то должно существовать знание о будущем, уже реализующиеся в форме футурологии, прогностики и других форм исследования будущего, которые могут быть включены в педагогический процесс как нечто темпорально-симметричное традиционному историческому подходу.

Без информатизации формирование такой качественно новой формы грядущего глобального образования в принципе невозможно и поэтому наиболее радикальные инновационные процессы, как можно ожидать, будут связаны с переходом к информационному обществу, являющемуся важной составляющей планетарного перехода на путь УР. Информационному обществу с УР, базирующемуся на знании и информации, должна соответствовать новая глобальная опережающая инновационно-образовательная система [17, с. 185].

Образование не сводится только к обучению и воспитанию в специальных образовательных учреждениях, а

видится как более широкий процесс информационного взаимодействия и в какой-то мере даже коэволюции человека и мира. И хотя дальнейшее существование и развитие индивида теперь уже требует непрерывного образования, особенно профессиональной подготовки, переподготовки и повышения квалификации, тем не менее «обучение жизнью» как естественный и непрерывный процесс пока играет не меньшую, а даже большую роль, поскольку в ряде областей деятельности человек в основном самообразовывается и образовывается жизнью, окружающей его социоприродной средой. «Образование» личности в таком случае происходит за счет адаптации к происходящему в жизни и окружающей среде.

Формирование человека осуществляется как в специальных образовательных учреждениях (системно-организованное образование), так и происходит «образовывание» человека самой жизнью (иногда оно именуется естественным образованием), причем весьма значительная часть населения планеты (около миллиарда человек) даже в настоящее время никогда не училась в каких-либо учебных заведениях. Впрочем, и та часть населения, которая посещает эти заведения, также испытывает стихийное влияние окружающей социальной и природной среды, «образовывается», взаимодействуя с социумом и природой. Предполагается, что в образовательных учреждениях человек получает основную часть своего «образовательного потенциала». Хотя никто серьезно не изучал, какая часть этого потенциала получена в процессе специального обучения, а какая в результате естественной жизнедеятельности человека. Именно эта последняя составляющая целиком определяла «образование» человека на начальных этапах человеческой истории.

В сфере образования взаимодействие человека и мира в информационном ракурсе сейчас наиболее отчетливо проявляется на пути становления так называемого «умного образования» (smart education). Интенсивно развивае-

мая в основном на Западе концепция Smart education, или умного образования [18], предполагает перенос образовательного процесса преимущественно в электронную окружающую среду, обеспечивающую эффективное обучение в этой интерактивной среде с помощью контента со всего мира, находящегося в свободном доступе (СМИ, Интернет, электронные библиотеки, социальные сети, неформальные сообщества и др.). Smart education за счет внедрения электронного взаимодействия, сделает обучение доступным везде и всегда, объединяя учебные заведения и профессорско-препо-давательский состав для реализации совместной образовательной деятельности в сети Интернет (являющейся одним из основных источников знания для студента). Smart education, или умное образование, - это гибкое обучение в интерактивной образовательной среде, осуществление образовательного процесса главным образом в сети Интернет на базе общих технологий, стандартов, соглашений.

Умное образование не только будет существенно стимулировать становление информационного общества во всех своих мыслимых вариантах (общества знания, общества образования и т.п.), но и приблизит его превращение в ноосферу. Умное образование становится одним из будущих фундаментов ноосферного образования и должно будет функционировать как компонент сферы разума. В определенном смысле умное образование в целом должно становиться опережающей и наиболее интеллектуальной социальной системой, которую можно будет считать совместно с образованием для УР в качестве начального этапа становления ноосферного образования.

В каком-то смысле образование – это не просто одна из четко выделенных форм человеческой деятельности, а вся совокупная человеческая деятельность, вся сфера информационного взаимодействия индивида с социальным и природным бытием (т.е. не только один из процессов его социализации), отображаемая в информационной эволю-

ции индивида. При таком широком видении образования оно оказывается элементом, или одним из множества глобальных процессов в потоке универсальной эволюции и может быть положено в основание наиболее общей концепции образования как информационного глобально-космического процесса.

И здесь мы попадаем в сферу так называемого онтологического подхода в педагогике, когда образование рассматривается не только как индивидуальная, общественная, но и всеобщая мироустроительная форма развития, когда эволюционирует весь сопряженный с человеком универсум – социокультурный, природный, космический, антропологический [19].

С позиций онтологического подхода образование должно быть направлено на формирование (образовывание) человека, что лишь частично учитывается в концепции образования как «передаче научных знаний и формировании минимально необходимого для включения в общественное производство набора умений и навыков умственной работы» [20, с. 169].

С этих наиболее общих позиций образование рассматривается не только как индивидуально-общественная (в системе «человек-социум»), но и как всеобщая миросозидательная форма развития, когда коэволюционирует весь сопряженный с человеком универсум – не только социокультурный, но и глобально-природный (а также вселенско-космический) и в информационном плане воздействует на формирование личности.

В этом, пожалуй, наиболее широком информационном представлении образовательного процесса формирование человека видится сквозь «призму» универсального (глобального) эволюционизма уже в более широкой – социоприродной системе «человек–общество–природа». Это не просто один из процессов социализации человека как его обычно представляют, а его восприятие информационного эволюционного взаимодействия со всей остальной

Вселенной в ее социальном и природном ракурсе. С позиций этого (назовем его также универсально-эволюционным) подхода образование реализует функцию формирования (образовывания) человека всеми информационными процессами в планетарно-космической системе эволюции «человек-общество-природа». Образование предстает в этой интерпретации не просто как процесс социализации индивида, а как эволюционно-информационный и даже информационно-коэволюционный процесс взаимодействия человека, общества и природы.

Поскольку эволюционный подход к образованию выходит за пределы педагогических наук, то речь, в частности, идет о генезисе того информационного процесса, который позже получил наименование обучения в мире живой материи. Педагогика вообще исследует образование лишь как развитие жизненного опыта человека [21, с. 25], т.е. в антропоцентрическо-эволюционном ракурсе. И иная постановка в традиционной педагогике представляется неприемлемой. Однако эволюционный антропоцентризм как некое новое видение глобально-универсального эволюционизма (или, по меньшей мере, одно из его направлений) захватывает многие науки, в том числе и космические (антропный космологический принцип). И важно видеть какое место здесь занимает обучение, а в перспективе и образование, принимающее уже не только глобальные, но и космические формы своей эволюции [22].

И все же при рассмотрении образования в широком глобально-эволюционном ракурсе оно «растворяется» и как-то теряется в других взаимодействиях человека и мира, общества и природы. Поэтому важно не только расширить понимание образования до глобально-космических масштабов, но и сузить до его понимания как всё-таки особого социального и социоприродного процесса. И это возможно, лишь выделив во всех мыслимых взаимодействиях его отражательно-информационный и коммуникативный аспект,

ориентированный на социализацию и развитие личности (а может быть, и некоторых социальных систем). Ведь образование, несмотря на то, что оно имеет свою материально-техническую и вещественно-энергетическую составляющие, в первую очередь является информационно-коммуникатив-ным процессом в системе «человек-общество-природа». И именно все формы движения информации оказываются главными среди других взаимодействий и их аспектов, как в узком понимании феномена образования, так и в широкой – глобально-эволюционной трактовке. Информационный аспект (составляющая) образования является основной его характеристикой, и это становится более очевидным при рассмотрении процессов футуризации образования и развития инновационно-опережающих процессов.

Представляя будущее глобальное образование как «образовывание», формирование человека не только в социально организованных, специальных учебных заведениях, но и всей жизнью человека, всей окружающей средой, уместно обратить внимание на два существенных обстоятельства. Во-первых, переходя от представления образования как трансляции знаний в специальных образовательных учреждениях («знаниевого» и «компетентностного» подхода») к более широкому «глобально-эволюционному» видению феномена образования, важно все же сконцентрировать внимание на том общем, что им присуще – движению и восприятию социально-значимой информации. В любом случае именно информационный аспект оказывается наиболее существенным и для «компетентостно-знаниевого» и для «глобально-эволюционного» видения образования. Формирование человека в процессе образования в обоих подходах может моделироваться процессом движения и накопления информации, т.е. мыслится как своего рода эволюционный информационный процесс, о котором здесь идет речь. Во-вторых, это процесс, носящий принципиально социоприродный характер, хотя обычно

его представляют в качестве чисто социального процесса, особенно, если речь идет об образовании как трансляции (устаревшего) знания. Переход от «трансляционно-знаниевого» к «информационно-эволюционному» видению обсуждаемого феномена позволяет представить образование в широком смысле как социоприродный процесс движения информации от общества и природы к человеку. В результате этого социоприродного процесса освоения информации человек увеличивает свое информационно-интеллектуальное содержание, он прогрессивно развивается и обретает способность самосохраняться в качестве самостоятельного и активного элемента общества и природы. Такое «социоприродно-информационное» видение образования соответствует как концепции глобального эволюционизма, так и социоприродным концепциям эволюции в форме УР и ноосферогенеза.

Только на пути информатизации и формирования виртуальной реальности в человеческой деятельности вообще, а в науке и в образовании в частности. появляется возможность опережающего моделирования реальности. Будущий переход к УР тем самым возможен лишь благодаря образованию в специальных учебных заведениях, а не в процессе естественной адаптации к окружающей действительности. Часть адаптивного потенциала необходимо ориентировать на будущую и пока виртуальную реальность новой модели, в которой предполагается жить грядущим поколениям. Желаемую и должную виртуальную реальность необходимо создавать уже сейчас, учитывая сильную инерционность и «эволюционную консервативность» процесса перехода от старой к новой модели развития.

Становление в будущем ноосферного интеллекта, наиболее полно реализующего опережающее моделирование социоприродных процессов, приведет к тому, что станет возможным не только отображение настоящего и прошлого, но и прогнозирование будущего и на этой базе

формирование рационального, эффективного и непрерывного социоэкоразвития.

Представляя образование как «образовывание», формирование человека не только в специальных учебных заведениях, но и всей жизнью, всей окружающей средой, уместно обратить внимание на два существенных обстоятельства. Во-первых, переходя от представления образования как трансляции знаний в специальных образовательных учреждениях («знаниевого подхода») к более широкому «онтологическому» видению образования, важно все же сконцентрировать внимание на том общем, что им присуще – движению и восприятию информации. В любом случае именно информационный аспект оказывается наиболее существенным для «знаниевого» и для «онтологического» видения образования. Формирование человека в процессе образования в обоих подходах может моделироваться процессом движения и накопления информации, т.е. мыслится как своего рода эволюционный процесс, о котором уже выше шла речь.

Во-вторых, это процесс, носящий принципиально социоприродный характер, хотя обычно его представляют в качестве чисто социального процесса, особенно, если речь идет об образовании как трансляции (устаревшего) знания. Переход от чисто «знаниевого» к информационно-онтологическому видению позволяет представить образование в том широком смысле как социоприродный процесс движения информации от общества и природы к человеку. В результате этого социоприродного процесса освоения информации человек накапливает свое информационное содержание, он прогрессивно развивается и вместе с тем обретает способность самосохраняться в качестве самостоятельного и активного элемента общества и природы. Такое «социоприродное информационное» понимание образования соответствует как концепции универсального эволюционизма, так и концепции УР.

Заключение

В дальнейшей эволюции образования на приоритетное место выходят информационные процессы и факторы, особенно новые информационные технологии и другие составляющие информатизации общества, включая искусственный интеллект, Интернет-медиатиза-ция и различные формы виртуальной реальности. Информационный аспект образовательной деятельности важен не только по формам и технологиям ее более эффективной реализации, но и по своему существу, ориентации инновационно-опережающих процессов в интересах реализации новой цивилизационной парадигмы.

Появление на информационной основе инновационно-опережающих процессов в образовании приведет к существенной его футуризации и становлению опережающего образования для УР и дальнейших эволюционных процессов цивилизации. «Устойчивая», а затем и ноосферная формы образования будут не только быстрее развиваться относительно других форм социальной деятельности, но и акцентировать внимание на будущем в самом содержании обучения, воспитания и творчества, исходить из принципа темпоральной целостности. В этом смысле произойдет сдвиг акцентов от модернизации образования к его футуризации и в перспективе – к становлению ноосферного образования через образование для УР и «умное образование».

Образование из современного механизма трансляции устаревших знаний, умений, навыков (и сопряженной с ними культуры) должно превратиться в приоритетный фактор социального развития в форме УР, обогащения новейших и прогностических знаний, умений, способствуя росту понимания, осознания происходящего и будущего. Образование тем самым должно превратиться из социального механизма, концентрирующего устаревшую научную

и иную социальную информацию для передачи ее последующим поколениям, в приоритетное средство выхода из системно-глобального кризиса цивилизации и выживания человечества через УР. Нынешняя же система образования (именно как система, а не отдельные ее элементы) не является таким средством, сколько бы финансов и усилий не направлялись в нее, в силу модернизационно-парадигмальной недостаточности.

Эволюционные трансформации мирового образования станут реализоваться вначале на пути преобразования современного образования в образование в интересах УР как основного образовательного процесса XXI века, содержание и развертывание которого направлено на реализацию целей и принципов новой модели развития мирового сообщества. В дальнейшем образование для УР в ходе развития «умного образования» и последующего ноосферогенеза постепенно превратится в ноосферное образование, а параллельно с этим, а тем более в перспективе – в «глобально-эволюционное» образование, ориентированное на существенно более развитый ноосферной наукой глобальный (универсальный) эволюционизм как основное интегративно-общенаучное ядро научно-образовательного процесса нынешнего столетия.

Выживанию человечества будут способствовать различные формы становящегося глобального образования, ориентированные на эффективное содействие выходу цивилизации из углубляющегося планетарного антропогенного кризиса. В результате постепенного перехода на упомянутые выше перспективные модели образования предполагается становление человека нового типа - ноосферной личности, подготовленного к реализации идеи выживания человечества, его дальнейшей непрерывной эволюции в сохранившейся биосфере и за ее пределами.

Список литературы

1. Урсул А. Д. Всеобщий характер информации // Современные проблемы математики, информатики и автоматизации: Материалы международного семинара. - Севастополь: Изд-во СевНТУ, 2012. - С. 5-20.
2. Сноу Ч. П. Две культуры. - М.: Прогресс, 1973.
3. Ильин И.В., Урсул А.Д., Урсул Т.А. Глобальный эволюционизм: Идеи, проблемы, гипотезы. - М.: Изд-во МГУ, 2012.
4. Ильинский И.М. Образовательная революция. - М.: Изд-во Московской социально-гуманитарной академии, 2002.
5. Урсул А.Д. Образовательная революция в перспективе устойчивого будущего // Знание. Понимание. Умение. 2009. № 1. - С. 15-18.
6. Эшби У.Р. Введение в кибернетику. - М.: Иностранная литература, 1950.
7. Ильинский И.М. Между прошлым и будущим. Социальная философия происходящего. - М.: Изд-во МосГУ, 2006.
8. Coombs PH**.** The World education crisis. A systems analysis. - Paris, 1968.
9. Coombs P.H. The World crisis in education: the view from the eighties. - New York, 1985.
10. Образование: сокрытое сокровище. - Париж: ЮНЕСКО, 1997.
11. Урсул А.Д. Информация и глобальные процессы: междисциплинарные исследования // Знание. Понимание. Умение. 2013.№ 3. - С.36 -42.
12. Итоговый документ Конференции ООН по устойчивому развитию «Будущее, которого мы хотим» - //www.un.org/ru/sustainablefuture.
13. Урсул А.Д. Футуризация как темпоральный феномен // NB: Философские исследования. 2013. № 7. - С.426-470.

14. Урсул А.Д. Принцип темпоральной целостности и образование // Alma mater/ Вестник высшей школы. 2008. № 3. - С. 28-35.
15. Урсул А.Д., Урсул Т.А. Эволюционные парадигмы и модели образования XXI века // NB: Педагогика и просвещение. 2012. № 1. - С.1-67.
16. Урсул А.Д. Контуры безопасно-устойчивого развития: размышления после саммита Рио+20 // Безопасность Евразии. 2012.№2. - С. 99-117.
17. Колин К.К. Актуальные философские проблемы информатики. Теоретические проблемы информатики. Т. 1. - М., 2009.
18. Тихомиров В.П. Смарт-образование – новые возможности для развития в информационном обществе // Информация и научное мировоззрение: новая миссия школьной библиотеки в XXI веке / Сб. статей. - М.: Русская школьная библиотечная ассоциация, 2013. - С. 303-319.
19. Пищулин Н.П. Философия образования: законы, доктрина, принципы. - М.: Изд-во МГПУ. 2001.
20. Куликовская А.В. Дидактика креативного автопоэзиса человека и его знания // Что такое постнеклассическое знание. - М.-Александров, 2007. - С. 169-178.
21. Новиков А.М. Основания педагогики. - М.: Эгвес, 2010.
22. Урсул А.Д., Урсул Т.А. Феномен образования в глобально-эволюционном ракурсе // Политика и общество. 2010. № 5. - С. 44-51.

УДК 004.492.3: 004.7
А.В.Скатков[1], д-р техн. наук, проф.; В.С.Ловягин[1], канд. техн. наук; С.А.Чорномыз[2]
[1]Севастопольский национальный технический университет, г.Севастополь, Украина
[2]Керченский колледж экономики и информационных технологий, г. Керчь, Украина
kvt.sevntu@gmail.com

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ КРИТЕРИАЛЬНОГО ОБНАРУЖЕНИЯ АТАК В ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ КРИТИЧЕСКОГО ПРИМЕНЕНИЯ

Основная тенденция развития современных телекоммуникаций — создание глобального информационного пространства и глобальных сетей. Развитие информационных технологий на базе телекоммуникационных сетей (ТКС) приводит к повышению эффективности промышленного производства и многих инфраструктур в энергетической, транспортной, социальной сферах. В большинстве случаев эти инфраструктуры критические (КИ). Этот факт накладывает безусловные ограничения на требования к гарантоспособности и безопасности критических ТКС (КТКС). КТКС, как одна из основных подсистем КИ, должна обеспечивать высокую реактивность, достоверность и своевременность передачи информации между функциональными подсистемами инфраструктуры. Отказ или нарушение работы КТКС может привести к дезорганизации функционирования всей инфраструктуры и, следовательно, опасности возникновения критических ситуаций, которые могут иметь катастрофические последствия [1]. Несмотря на интенсивное развитие компьютерных систем и информационных технологий защиты, КТКС продолжа-

ют быть уязвимыми для внешних и внутренних вирусных атак (ВА). Основной угрозой безопасности КТКС являются акты несанкционированного перехвата каналов обмена информацией и управления. Несанкционированный перехват каналов может быть осуществлен в результате нарушения прав доступа (НПД) субъектов к объектам КТКС. Далее будем называть ВА и акты НПД А-событиями.

При наличии атак и актов нарушения прав доступа возникает существенное изменение системных показателей КТКС: одним из этих показателей является относительная функциональная быстродействие процессорных каналов. Процесс изменения системных показателей КТКС предлагается характеризовать посредством четырех непересекающихся зон (A, B, C, D) и фазами *(τi; τj)* (рисунок 1).

На рисунке 1: *(τ1; τ2*) — фаза действия А-события, а именно, развития ее последствий в КТКС до момента ее обнаружения, при этом функциональная быстродействие снижается с *v1* в *v2*; *τ2* — момент обнаружения А-события; *(τ2; τ3)* — фаза принятия контрмер в А-события (прекращение развития последствий ее действия), величина *(v2 — v3)* — системные ресурсы, привлекаемые для осуществления контрмер; *(τ3; τ4)* — фаза восстановления системы, то есть устранение последствий А-событий.

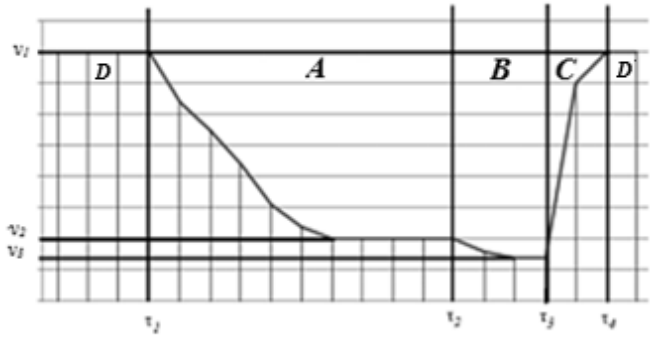


Рисунок 1 — Изменение относительной производительности КТКС в условиях наличия А-события (зона А), противодействия ей (зона В) и восстановления (зона С).

Размер площадей зон A и B пропорционален объему потерь до и после обнаружения А-события соответственно. Размер площади зоны C определяет эффективность восстановления, связанного с повышением быстродействия с ν3 к ν1. Зона D — нормальное функционирование КТКС. Установлено, что качественный вид изменения функциональных характеристик сохраняется при различных типах А-событий, но отличается параметрами.

Анализ последствий действий основных видов А-событий, современных средств и методов их обнаружения в КТКС [2] выявил ряд недостатков существующих детекторов А-событий в КТКС. Одним из существенных недостатков является отсутствие комплексных методологических решений задачи своевременного обнаружения А-событий в условиях информационного и временного дефицита, что особенно важно для КТКС.

Результаты анализа существующих методов оценки системных характеристик КТКС в условиях деградации и восстановления показали, что в них не учтена многоэтапность таких процессов, затрудненное прогнозирования оценок системных характеристик. Присутствует конфликтная ситуация, невозможность повышения быстродействия процессов обнаружения А-событий, наряду с повышением достоверности. Задача требует решений векторной оптимизации в рамках компромиссов, Парето–решений. необходимости минимизации реактивности обнаружения А-события и обеспечения максимизации достоверности их обнаружения. С учетом этого сформулирована основная системная задача – разработка методов и информационных технологий минимизации этапов $(\tau_1;\tau_2)$ за счет своевременного обнаружения А-событий и $(\tau_3;\tau_4)$ за счет информационного обеспечения системы поддержки принятия решений для динамического распределения ресурсов.

Системный анализ задачи выбора параметров эффективного КДА и вариантный анализ выбора среды разработки программной реализации информационной техно-

логии (ИТ) выбора оптимальных параметров КДА позволили предложить структуру и на этой основе реализовать ИТ выбора оптимальных параметров КДА.

В основу реализации выбранного направления исследований легло представление КТКС кортежем (1).

$$S = \langle \Phi, X, Z, U, K_a, P \rangle, \quad (1)$$

где Ф = {φi} — множество функций КТКС; X = {Xf} — множество входных пользовательских пакетов задач; Z = {Zk} — множество пакетов результатов выполнения пользовательских задач; U = {Ul} — множество операторов выбора выполняемых функциональных преобразований, необходимых для выполнения пакета пользовательских задач; Кеф — функционал оценки эффективности работы КТКС, Р — правила поддержки принятия решений по управлению стратегиями обнаружения А-событий. Для обеспечения эффективного функционирования ТКС необходимо решить оптимизационную задачу (2).

$$M\left[K_{эф}\left(R, \Lambda, S, C_s, K_k\right)\right] \rightarrow extr, \quad (2)$$

где $S=\{S_i\}$ – множество стратегий обнаружения А-событий в КТКС; $R=\langle R_i \rangle$ – вектор системных ресурсов, выделяемые на противодействие А-событиям в КТКС; $\Lambda=\langle \lambda_a \rangle$ – вектор интенсивностей атак в КТКС; C_s – вектор системных характеристик КТКС; $K_к$ – функционал эффективности применяемых к А-событиям контрмер. Решение оптимизационной задачи (2), при вариации вектора R, возможно минимизировать фазу $(\tau_3;\tau_4)$. Для обеспечения эффективного функционирования системы защиты КТКС, необходимо решение оптимизационной задачи (3).

$$M\left[K_k\left(R, \Lambda, S\right)\right] \rightarrow extr, \quad (3)$$

где $S=\langle n, C_{об}, \alpha, M \rangle$, n — объем выборок анализируемых функциональных характеристик КТКС, Соб — критерий выявления А-событий; α — нормативный уровень достоверности, М — вычислительная мощность критерия. На

основании решений (3) при вариации вектора S = <n,Соб,α,M>, возможно минимизировать фазу (τ1; τ2).

Предлагается принцип функционирования КДА. Даны пары таблиц (x_i,y_i), i=[1,n], где x_i и y_i – нормативное и текущее значения функциональных характеристик КТКС соответственно. Полагаем, что $z_i=y_i-x_i=\theta+e_i$, где e_i – независимая случайная величина с медианой Med(e)=0. Формулируются гипотезы H_0: «А-событие статистически не различимо в КТКС». H_0 принимается при $\theta=\theta_K$, где θ_K – чувствительность критерия. Односторонняя альтернатива H_1 принимается при $\theta>\theta_K$. КДА описывается как (4).

$$D=\langle X,Y,\theta_k,H_0,H_1,\alpha,\beta,Mk,Rk,N\rangle \qquad (4)$$

где Mk — множество непараметрических критериев, Rk — правила принятия решений на основе Mk, N — объем выборочных данных; α, β — вероятности допущения ошибок первого и второго родов соответственно.

Аналитические выражения для (1) - (4) в большинстве случаев неизвестны, и нахождение решений оптимизационных задач методом полного перебора в данных условиях практически невозможно, поскольку это NP-полная задача. С целью решения для практически значимых случаев предложено использование информационной технологии критериального обнаружения атак (ИТ КВА), которая содержит настраиваемый задачей компонентов вектора S критериальных детектор атак (КДА) и блок оценки эффективности принимаемых контрмер для решения задачи динамического распределения ресурсов на восстановление КТКС.

Необходимость построения ИТ обуславливается невозможностью аналитического определения мощности непараметрических статистических критериев в данных условиях. ИТ реализована на основе специального программного комплекса для исследования мощности и применимости статистических непараметрических критериев «ИСМК». «ИСМК» является основой системы ППР по на-

стройке КДА (СППР-ВС), т.е. поиска решений (3). На рисунке 3 приведены основные функции, компоненты и структура «ИСМК».

Программный комплекс «ИСМК» содержит в своем составе следующие основные модули и блоки: блок управления экспериментальным исследованием, позволяющий в диалоговом режиме управлять статистическим экспериментом; база данных статистических критериев; подсистема арифметико-аналитического вывода; подсистема статистического анализа; блок сравнения критериев согласно выбранной стратегии сравнения.

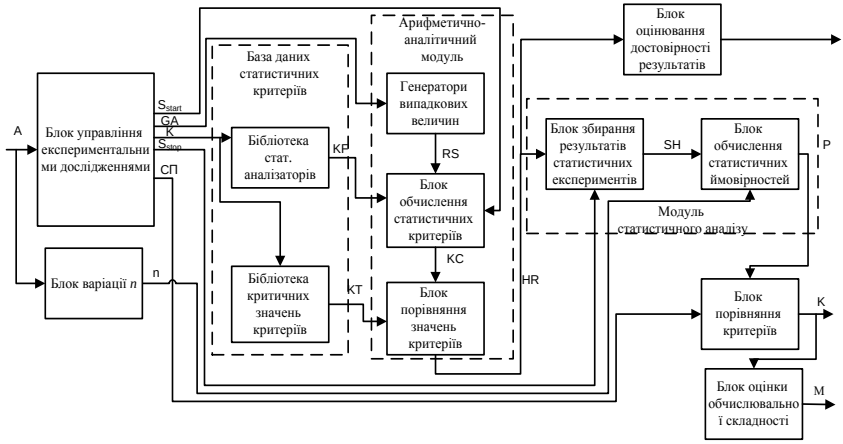


Рисунок 3 — Структура «ИСМК» для исследования мощности непараметрических критериев и построения областей эффективного применения для задачи обнаружения A-событий в КТКС.

A, GA – параметры управляемого вероятностного моделирования и генераторов случайных величин; Sstart, Sstop – команды начала и окончания статистического эксперимента, соответственно; K – вектор критериев для анализа; n – количество произведенных параллельных итераций эксперимента; СП – параметры выбора оптимального критерия, $СП=\{\pi,\alpha,\beta,P(H0|H0),P(H1|H1)\}$; π – мощность критерия; α, β – вероятности допущения ошибок первого и второго рода, соответственно; P(H0|H0), P(H1|H1) – веро-

ятности принятия гипотез; RS – выборки случайных величин; *KF* – функции, реализующие вычисление критериев; *KC, KT* – вычисленные и табличные значения исследуемых критериев, соответственно; *HR* – принятие или отвержение гипотез H0 или H1 для исследуемых критериев; *SH* – статистика принятия гипотез; *P* – вектор вероятностей *P(H0|H0), P(H1|H1), P(H0|H1), P(H1|H0)*; K – выбранный критерий.

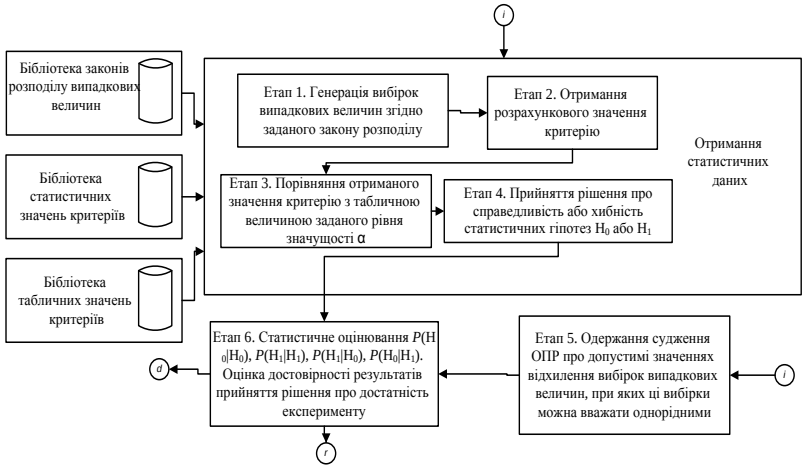
На основе анализ областей решений (3) найдена точка, доставляющая экстремум, однако в случае действия неучтенных факторов, дрейфа случайных параметров, возникает задача построения областей допустимых значений решений (3). Процесс построения областей эффективной применимости исследования представляет собой вероятностное моделирование, позволяющее получить оценки рисков при различных сочетаниях входных параметров. Базовый сценарий эксперимента предоставлен на рисунке 4, где *i* – альтернативные выборы ЛПР; *r* – выходные данные для СППР по управлению системой защиты (СППР-К) ТКС; *d* – оценка достоверности и достаточности эксперимента.

Управляющие альтернативные выборы ЛПР *i* описываются кортежем (5).

$$i = \left\langle K, C_{\Pi}, \alpha, |RS|, N, DRS, DRS_{o}, \xi_1 \right\rangle, \qquad (5)$$

где *DRS* – границы изменения параметров распределений исследуемых выборок функциональных характеристик КТКС, DRS_o – допустимые значения изменения параметров выборок функциональных характеристик КТКС.

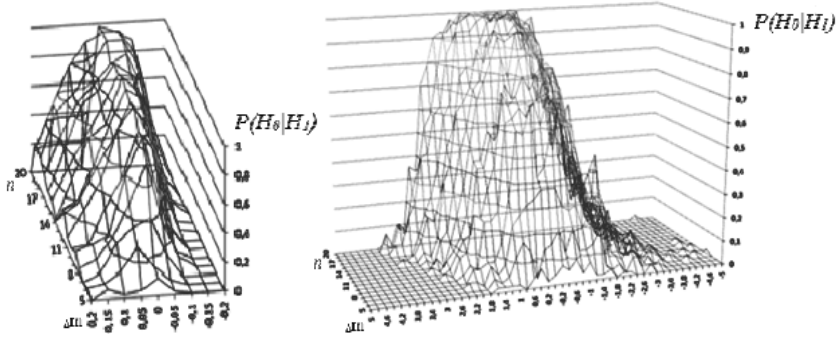
Выходные данные для СППР-НД *r* могут быть описаны как: $r=<K_q,R_q>$, где $q=\{1;|K|\}$; R – ранг K_q критерия, согласно СП.



Р

исунок 4 — Схема типичного сценария вероятностного моделирования для построения областей эффективных решений задачи (3) непараметрических критериев.

На основе комплексных экспериментальных исследований построен атлас зависимостей вероятностей $P(H_0|H_0)$, $P(H_0|H_1)$, $P(H_1|H_0)$, $P(H_1|H_1)$ от $_{\Delta}RS$ для детерминированного подхода и в условиях нечеткой логики, что позволило оценить риски. Пример найденных зависимостей приведен на рисунке 5.



а) б)

Рисунок 5 — Пример зависимостей Р (Н0 | Н1) от степени смещения для нормальных (а) и равномерных (б) функций распределения сравниваемых выборок и их объемов.

На основании результатов статистического моделирования рассчитаны риски как произведение (6).

$$R_{\Delta m,n} = C \cdot P(H_0 \mid H_1)_{\Delta RS,|RS|}, \quad (6)$$

где $C=f(\Delta RS,|RS|)$ – величины потерь связанных с ошибками первого рода; $P(H_0|H_1)_{\Delta m,n}$ – вероятности допущения ошибки первого рода.

Декартово произведение векторов C и $P(H_0|H_1)_{\Delta m,n}$ позволило получить оценочную матрицу выбора стратегии управления рисками при обнаружении А-событий в КТКС. Тем самым решена задача повышения адекватности систем поддержки принятия решений по управлению рисками при принятии решений по противодействию А-событиям (СППР-УР) за счет решения задачи исследования мощности и областей эффективной применения КДА для детерминированных данных и в условиях нечеткости.

Наличие А-событий в КТКС, несвоевременное их обнаружение, недостаточность мероприятий по компенсации последствий А-событий приводит к деградации функциональных характеристик ТКС, вплоть до перехода системы в состояние отказа в обслуживании (DoS). С целью недопущения перехода КТКС в состояние DoS за счет своевременного обнаружения и привлечения достаточного объема ресурсов в процессе восстановления КТКС, необходимо решение оптимизационной задачи динамического распределения ресурсов.

С этой целью разработана аналитическая модель в классе полумарковских процессов (ПМП), позволяющая провести параметрическую идентификацию системных характеристики КТКС, в частности время до попадания системы в состояние DoS отказа.

Функциональные характеристики КТКС деградируют в условиях наличия А-событий при несвоевременном обнаружении или непринятии достаточных мероприятий по компенсации последствий А-событий. Множество точек N задает последовательность этапов деградации функцио-

нальных характеристик КТКС при возникновении А-события. Состояние с номером N – состояние DoS. Состояние с номером 1 – функционирование КТКС в области *D.*

Переход в следующее *(i+1)* состояние (деградация) происходит в случае, если время η_i, которое система может функционировать до принятия достаточных мер по компенсации А-событий, превысило h_i – момент окончания выполнения работ по восстановлению КТКС после А-событий. Определение перехода на следующее состояние или восстановление в предыдущее осуществляется на основании настраиваемого КДА, путем проверки гипотез H_0, H_{11}, H_{12}, где H_{11} свидетельствует о деградации функциональных характеристик КТКС, а H_{12} о их восстановлении. Величина η_i – случайная величина с функцией распределения $F_i(x)$ и математическим ожиданием $M\{\eta_i\}$. $(\tau_3;\tau_4)$ – время до перехода в S_1, т.е восстановления системы за счет проведения мероприятий по устранению последствий атак

Структура переходов по состояниям вложенной цепи Маркова для процесса деградации функциональных характеристик КТКС в условиях воздействия А-событий при частичных восстановлениях предоставлена на рисунке 6, где $\xi_{i,j}$ – вероятность перехода из состояния *i* в состояние *j*.

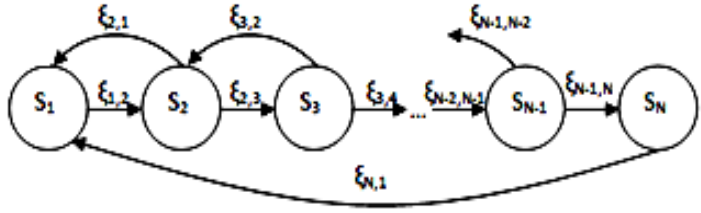


Рисунок 6 — Вложенная марковская цепь для процесса деградации и частичного восстановления функциональных характеристик КТКС в условиях действия А-событий.

Аналитическое выражение оценочного функционала для нахождения математического ожидания времени перехода КТКС в состояние DoS в условиях наличия А-событий может быть описано как (7).

$$t_{1,N}=\sum_{i=1}^{N-1}\frac{b_i}{\prod_{k=i}^{N-1}F_k}-\sum_{i=1}^{N-2}\frac{b_i}{\prod_{k=i}^{N-1}F_k}\sum_{l=i+1}^{N-1}F_l\overline{F_{l+1}}+\sum_{i=1}^{N-4}\frac{b_i}{\prod_{k=i}^{N-1}F_k}\sum_{l=i+1}^{N-3}F_l\overline{F_{l+1}}\sum_{m=l+2}^{N-1}F_m\overline{F_{m+1}}-\sum_{i=1}^{N-6}\frac{b_i}{\prod_{k=i}^{N-1}F_k}\sum_{l=i+1}^{N-5}F_l\overline{F_{l+1}}\sum_{m=l+2}^{N-3}F_m\overline{F_{m+1}}\sum_{j=m+2}^{N-1}F_j\overline{F_{j+1}}+\ldots$$

где $b_i=h_i-F_i(h_i)(h_i-M\{\eta_i\});F_i=F_i(h_i)$

Аналогичным образом была предложена модель многоэтапной деградации КТКС с полным восстановлением (рис. 7).

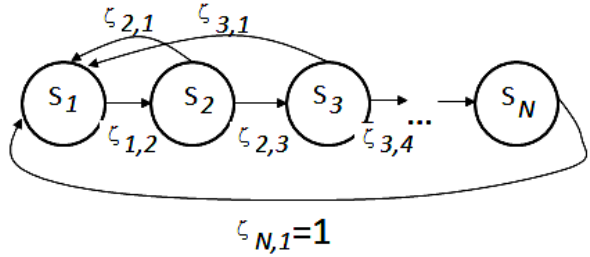


Рисунок 7 — Вложенная марковская цепь для процесса деградации и полного восстановления функциональных характеристик КТКС

Методом полной математической индукции получено аналитическое выражение оценочного функционала для вычисления математического ожидания времени до перехода КТКС в состояние DoS (8)

$$t_{1,N}=\sum_{i=1}^{N-1}\frac{h_i-F_i(h_i)[h_i-M\{\eta_i\}]}{\prod_{m=1}^{N-1}F_m(h_m)} \qquad (8)$$

Параметрическая идентификация системных характеристик на основе (7) и (8) позволила повысить эффективность СППР для принятия контрмер в А-событий и проведения мероприятий по восстановлению КТКС за счет нахождения базовой оценки эффективности этих мероприятий.

Аналитические модели (7) и (8) позволяют получить точечную оценку времени до перехода КТКС в состояние DoS, однако нет возможности получить оценки дисперсии. С целью развития моделей (7) и (8) предложено имитаци-

онное моделирование процесса восстановления ТКС, которое учитывает особенности функционирования КТКС и позволяет производить динамическую оценку системных характеристик КТКС, в условиях действия А-событий, принятия контрмер к ним и проведения мероприятий по устранению их последствий.

Предложенная ИТ «ИС–Эф» позволяет в диалоговом режиме проводить динамическую оценку $t_{1,N}$, путем вариации h и $F(h)$. Пользовательский интерфейс (ПИ) разработанной ИТ «ИС–Эф» имитационного моделирования имеет иерархическую структуру. На первом иерархическом уровне ПИ позволяет оператору устанавливать параметры этапов деградации КТКС, получать траекторию деградации и восстановления, основную системную характеристику процесса – математическое ожидание времени до попадания КТКС в состояние DoS. Имитационная модель может быть функционально расширенно за счет добавления новых модулей.

На основе предложенной имитационной модели решается оптимизационная задача динамического распределение ресурсов, необходимых для проведения мероприятий по противодействию А-событиям и восстановлению системы после успешных атак. На основании решения данной задачи может быть построена высокоадекватная СППР по выбору стратегии реинженеринга КТКС (СППР-Ре), что позволит повысить функциональное быстродействие КТКС, гарантоспособность, безопасность, обеспечит требуемый уровень функциональной готовности.Интеграция ИТ «ИСМК», «ИС-Эф», КДА, СППР-К,-Р,-Ре позволяет получить ИТ КОА.

Направление дальнейших исследований состоит в построении коллектива решающих правил по обнаружению вирусных атак, а также исследование уровня робастности семейства критериальных детекторов.

Список литературы

1. Технологии высокой готовности для программно-технических комплексов космических систем / В.С. Харченко, О.Н. Одарущенко, Ю.Л. Поночовый и др. – Х.:Гос. Центр регулирования качества поставок и услуг, Нац. аэрокосм. ун-т им. Н.Е. Жуковского «ХАИ», 2010. – 372 с.
2. Девянин П.Н. Модели безопасности компьютерных систем: Учеб. пособие для студ. высш. учеб. заведений / П.Н. Девянин – М.: Издательский центр «Академия», 2005. – 144 с.

УДК 004.56
А.А. Брюховецкий, канд. техн. наук, доц.
А.В. Скатков, д-р техн. наук, проф.
П.О. Березенко, магистр
Севастопольский национальный технический университет, г. Севастополь, Украина
a.alexir@mail.ru

АДАПТИВНАЯ МОДЕЛЬ ОБНАРУЖЕНИЯ УЯЗВИМОСТЕЙ В КРИТИЧЕСКИХ ПРИЛОЖЕНИЯХ НА ОСНОВЕ РЕШАЮЩИХ ДЕРЕВЬЕВ И БАЙЕСОВСКОГО КЛАССИФИКАТОРА

В связи с интенсивным ростом информационных технологий и их внедрением в различные отрасли деятельности человека вопрос информационной безопасности становится все более актуальным. На сегодняшний день существует большое количество систем обнаружения уязвимостей достаточно эффективно выполняющих детализированные исследования и поиск злоумышленников на основе известных сигнатур атак. К недостаткам таких систем следует отнести: неспособность обнаруживать и блокировать неизвестные уязвимости, невозможность автоматического ввода новых контролируемых шаблонов, отсутствие возможностей прогнозирования действий злоумышленника и другие. Поэтому разработка адаптивной модели обнаружения уязвимостей представляет собой актуальную задачу[1-4].

В зависимости от источника обнаружения вторжений различают системы уровня host-based (HBIDS), сетевые СОВ (NIDS – network intrusion detection), а также системы уровня приложений (APIDS) и гибридные СОВ (HIDS), которые сочетают комбинированные методы [1]. Первые идентифицируют вторжения, анализируя события и трафик, поступающий на отдельный компьютер, в то

время как вторые – исследуют сетевой трафик. Системы уровня приложений , как правило, располагаются между web-сервером и, например, SQL-сервером. На рисунке 1 представлена структура компьютерной сети с альтернативным расположением СОВ. В свою очередь, системы обнаружения вторжений, в зависимости от используемой технологии выявления атак, разделяют на два основных типа: системы обнаружения злоумышленного поведения и системы обнаружения аномального поведения.

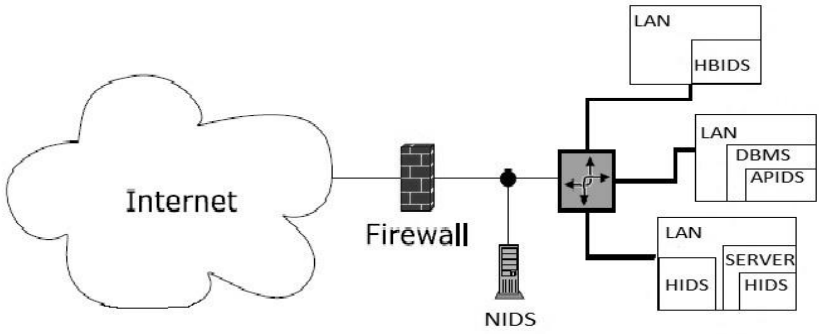


Рисунок 1 – Структура компьютерной сети с альтернативным расположением СОВ

Первые ориентируются на модель злонамеренного поведения (например, шаблон / сигнатура атаки) и сравнивают модель с потоком событий. На основании сравнительного анализа система принимает решение блокировать тот или иной пакет, либо пропускать для взаимодействия непосредственно с операционной системой. Сигнатурные системы обнаружения вторжений обладают высокой производительностью, эффективностью обнаружения при сравнительно невысоких требованиях к аппаратному обеспечению. Однако они имеют ряд недостатков: невозможность автоматического ввода новых сигнатур, отсутствие системы блокирования неизвестных сигнатур, отсутствие возможностей прогнозирования действий злоумышленника, отсутствие подсистемы мониторинга аппаратных ре-

сурсов и другие. Существующие сигнатурные системы обнаружения вторжений не могут профилировать перехваченный поток данных распределенных вторжений для их классификации и выработки сигнала о вторжении. Поэтому системы обнаружения вторжений второго типа используют методы для распознавания неизвестных атак. Такие СОВ проектируются, как правило, на основе моделей нормального поведения (например, профиль системы) и ищут аномальные вхождения в поток событий.

Задача исследования нормального и аномального поведения сетевых компьютерных систем является очень сложной и комплексной. Для ее решения используются, как правило, различные статистические модели для оценки вероятности появления заданных значений (событий). Низкая и высокая вероятность в этом случае может означать появление аномалий.

Для сравнительного анализа методов обнаружения вторжений могут быть выбраны следующие критерии:

- уровень наблюдения за системой;
- верифицируемость метода (позволяет провести экспертную оценку корректности метода и его реализации в произвольный момент времени, в том числе в процессе эксплуатации системы обнаружения на его основе);
- адаптивность метода (устойчивость метода к малым изменениям реализации атаки, которые не изменяют результат атаки);
- точность обнаружения вторжений и уровень ложных тревог и др.

Несмотря на широкое распространение СОВ и активные исследования в данной области, существует потребность в разработке новых методов и моделей обнаружения сетевых атак.

Целью работы является разработка модели обнаружения вредоносных программ по значениям признаков в сетевом трафике на основе решающих деревьев и модели байесовского классификатора. Разрабатываемая система

относится к СОВ, которые обнаруживают аномальное состояние трафика, а значит применяемая в ней технология позволяет обнаруживать новые виды атак.

Постановка задачи

Набор данных, который используется в работе для построения модели, был смоделирован и получен в компьютерной сети ВВС США с целью имитации компьютерных атак и содержится в базе данных лаборатории MIT Lincoln Labs [2]. Данные содержат около 4 Гб сжатого TCP сетевого трафика, который собирался 7 недель и составил около 5 миллионов записей о соединениях. Каждое TCP / IP соединение состоит из 41 количественного и качественного признака. Последний 42-ой признак характеризует состояние трафика. Оно может быть нормальным или аномальным. Каждая запись содержит около 100 байт. Признаки могут быть разделены на три группы:

- основные характеристики соединения, например, продолжительность соединения, тип протокола, сервис, число переданных байт от источника к приемнику и в обратном направлении, отдельные флаги. Некоторые значения признаков определяются с задержкой в течение определенного временного интервала;
- статистические характеристики трафика, которые вычисляются с использованием 2-х секундного временного окна или в течении большего временного промежутка. Характеристики подразделяются на две группы: атрибуты относящиеся к конкретному host – компьютеру или к конкретному сервису. Отдельные атаки сканирования портов выполняются дольше, чем 2 сек. Поэтому ряд признаков обрабатывается окном в 100 соединений;
- признаки внутри отдельного соединения. В отличие от большинства DoS – атак и сканирования портов, R2L и U2R атаки характеризуются отдельными непродолжительными проявлениями к отдельному компьютеру. В то время как DoS – атаки и Probing инициируют множест-

венные соединения в короткий промежуток времени.

Более подробная информация о признаках сетевого трафика может быть получена в [2]. На основе приведенных данных требуется построить дерево классификации.

Метод решения задачи

Метод решения базируется на оценке априорной вероятности о значениях признаков и апостериорной вероятности принадлежности векторов заданным классам. Алгоритм оценивает вероятность появления каждого класса векторов в исходном наборе данных, а также условную вероятность принадлежности каждого значения атрибута вектора к каждому из пяти классов. После этого алгоритм использует эти вероятности для обновления значений C_j для каждого вектора в наборе данных. Эта процедура выполняется на основе модели наивного байесовского классификатора. После этого вектору присваивается метка класса, соответствующая максимальному значению апостериорной вероятности. Затем алгоритм пересчитывает априорные вероятности появления классов и условные вероятности принадлежности значений атрибутов заданному классу. Повторно проводится классификация каждого вектора. Если встречаются ошибочные классификации, то исходное множество векторов разбивается на подмножества и формируется дерево решений.

Алгоритм исследует значения признаков каждого атрибута, которые предварительно упорядочиваются. Основная задача – определить значение признака, по которому будет происходить разбиение исходного множества.

Пусть задано множество примеров T, где каждый элемент этого множества описывается m атрибутами. Количество примеров в множестве T будем называть мощностью этого множества и будем обозначать |T|. Метка класса C принимает следующие значения C1, C2, …, Ck. Пусть

$freq(C_j,T)$ – количество примеров из множества T, относящихся к одному и тому же классу C_j. Тогда выражение

$$Info(T)=-\sum_{j=1}^{k}\frac{freq(C_j,T)}{|T|}\log_2\left(\frac{freq(C_j,T)}{|T|}\right) \quad (1)$$

дает оценку среднего количества информации. В терминологии теории информации данное выражение называется энтропией множества T. Ту же оценку, но только уже после разбиения множества T по атрибуту X на n подмножеств, дает следующее выражение (2):

$$Info_x(T) = \sum_{i=1}^{n}\frac{|T_i|}{|T|}Info(T_i) \quad (2)$$

Критерий оценки значения энтропии до и после разбиения – прирост информации, рассчитывается как (3):

$$Gain(X) = Info(T) - Info_x(T) \quad (3)$$

Далее вводится нормализация по формуле (4):

$$Split_Info_x(T) = -\sum_{i=1}^{n}\frac{T_i}{T}\log_2\left(\frac{T_i}{T}\right) \quad (4)$$

Выражение (4) оценивает потенциальную информацию, получаемую при разбиении множества T на n подмножеств. Основная задача – определить значение признака, по которому будет происходить разбиение исходного множества. Для этого используется критерий Gain_Ratio (5):

$$Gain_Ratio(X) = Gain(X) / Split_Info_x(T) \quad (5)$$

После разбиения исходного множества *T* на подмножества, процедура оценки вероятностей приведенная выше, повторяется для каждого из подмножеств. Выполнение алгоритма заканчивается в случае правильной классификации векторов или при выполнении условия, когда листья дерева будут содержать вектора одного класса.

В режиме контроля алгоритм осуществляет 2-х уровневую классификацию (Рисунок 2). На первом этапе

распознается два состояния трафика: нормальное или аномальное. На втором – уточняется класс вторжения.

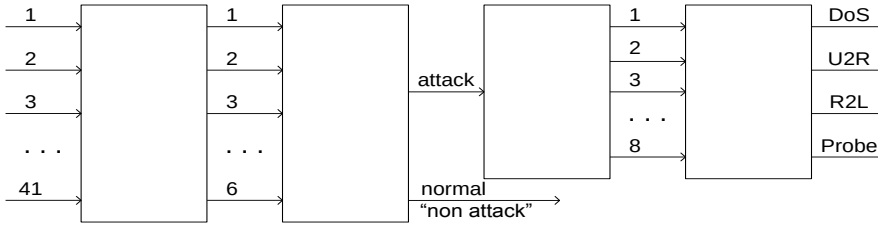


Рисунок 2 – Структурная схема классификации вторжений

Результаты экспериментальных исследований

Модель системы обнаружения вторжения реализована в программной среде – Eclipse. В результате исследования программной системы были получены основные характеристики. Эксперименты проводились на смешанной (нормальной и аномальной) обучающей и тестовой выборках данных базы KDD. Тестовая выборка проверялась по 14 числовым признакам сетевого трафика, шесть из которых использовались для классификации нормального и аномального трафика, и восемь – для определения класса вторжения. Эти признаки были определены как оптимальные после обучения модели.

В таблице 1 приведено количество записей обучающей и тестовой выборок, принадлежащих конкретным классам.

Таблица 1 - Число обучающих и тестовых записей

Классы данных	Обучающая выборка	Тестирующая выборка
Normal	107273	90592
DoS	441458	572594
R2L	1126	7683
U2R	52	70
Probing	3157	4189
Всего	553006	675128

Характеристики точности и уровня ложных срабатываний представлены в таблице 2. Выражения для определения этих характеристик приведены ниже [3, 4]:
DR – частота выявления атаки:

$$DR= TP/(TP+FN) \quad (6)$$

и FAR – уровень ложной тревоги (False Alarm,):

$$FAR=FP/(FP+TN) \quad (7)$$

где TP (true positive) - количество правильно классифицированных аномалий, FN (false negative) - количество образцов нормального трафика определенного системой как аномальный, FP (false positive) - количество аномальных образцов сетевого трафика определенных системой как нормальные, TN (true negative) - количество правильно идентифицированных системой образцов нормального трафика.

Указанные параметры позволяют оценить число правильно обнаруженных нормальных и аномальных образцов, а также - ошибки первого и второго рода.

Таблица 2 - Характеристики СОВ

	Normal	Probe	DOS	U2R	R2L
DR(%)	98,03	92,75	97,76	68,23	79,21
FAR(%)	0,3	5,43	1,27	7,32	9,41

Заключение

Предложена модель обнаружения вредоносных программ на основе решающих деревьев и модели наивного байесовского классификатора. Модель ориентирована на исследования значений признаков сетевого трафика. Разрабатываемая система контролирует аномалии в сетевом трафике. Применяемая в ней адаптивная технология позволяет обнаруживать новые виды атак. Основное внимание при разработке модели уделено выявлению наиболее информативных признаков сетевого трафика. Это позволяет существенно снизить размерность анализируемого пространства признаков, увеличить быстродействие, повысить точность обнаружения и снизить уровень ложных срабатываний.

Список литературы

1. Искусственные иммунные системы и их применение / Под ред. Д. Дасгупты. Пер. с англ. - М.: Физматлит, 2006. - 344 с.
2. KDD cup 99 Intrusion detection data set [Электронный ресурс]. – Электрон. текстовые данные (752 Мб). – Darpa: Irvine, CA 92697-3425, 1999. – /http://kdd.ics.uci.edu/databases/kddcup99/kddcup/task.html Monday, 17 March 2013 19:07:34.
3. Dasgupta D, Yu S., Majumdar N. MILA - Multilevel Immune Learning Algorithm / D. Dasgupta, S. Yu, and N. Majumdar // Proceedings of the Genetic and Evolutionary Computation Conference - 2003, Springer - Verlag: Berlin Heidelberg, 2003. - P. 183–194.
4. Ji Z., Dasgupta D. Real-valued negative selection algorithm with variable-sized detectors / Z.Ji, D. Dasgupta // Proceedings of the Genetic and Evolutionary Computation - 2004, Springer - Verlag: Seattle, WA, USA ,2004. - P. 287–298.

УДК: 004.8+159.95+612.8+530.145
А.Д. Панов, канд. физ.-мат. наук, ст. научн. сотр.
НИИ ядерной физики им. Д.В. Скобельцына МГУ им. М.В. Ломоносова, г. Москва, Россия
panov@dec1.sinp.msu.ru

ТЕХНОЛОГИЧЕСКАЯ СИНГУЛЯРНОСТЬ, ТЕОРЕМА ПЕНРОУЗА ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ И КВАНТОВАЯ ПРИРОДА СОЗНАНИЯ

Введение. Искусственный интеллект и технологическая сингулярность

Проблема искусственного интеллекта (ИИ) оказалась в центре внимания общества вместе с возникновением первых компьютеров в начале пятидесятых годов прошлого века, продолжает оставаться там сейчас, и ничто не предвещает изменение этого положения в любом обозримом будущем. Отношение к перспективам создания ИИ было разным. С одной стороны, всегда имелось достаточное количество оптимистов, которые считали, что создание ИИ является чисто технической задачей, решение которой будет обеспечено ростом мощности вычислительной техники. С другой стороны, многим постепенно стало понятно, что создание ИИ не является чисто технической задачей, но является чрезвычайно сложной междисциплинарной проблемой, затрагивающей в высшей степени фундаментальные проблемы бытия. К этой последней партии принадлежит и автор настоящей статьи, по мнению которого мы не приблизились не только к решению задачи создания настоящего сильного ИИ, но даже к внятной формулировке проблемы, которую хотим решить.

В настоящей статье мы постараемся отчасти закрыть эту брешь, подвергнув критике с единой (авторской) точки зрения оба крайних направления. Анализ концепции

технологической сингулярности (ТС) будет основан в основном на книге Рэя Курцвейла [1], анализ представлений Роджера Пенроуза на его книгах [2, 3]. Забегая вперед отметим, что точка зрения сторонников ТС, по нашему мнению, страдает чрезмерным оптимизмом, в то время как и выводы Пенроуза представляются излишне пессимистическими. Начнем с определений, или, точнее, с неформального разъяснения смысла используемых терминов. Под системами искусственного интеллекта будут пониматься автономные искусственные устройства, способные выполнять интеллектуальные функции. Типичным примером инструментальных систем является системы автоматического проектирования (САПР). Надо также отметить, что резкой границы между искусственным и инструментальным интеллектом провести невозможно. Под сильным искусственным интеллектом в этой статье будет пониматься такой ИИ, который превосходит человека, или по крайней мере не уступает ему, по всем интеллектуальным функциям, во всех отношениях. Заметим, что используемое здесь понятие сильного ИИ хотя и похоже на распространенное определение [4] или на оригинальное определение Джона Сёрла [5], но является более узким и более сильным, чем это обычно принимается. Такое понятие более адекватно представлению о ТС.

Предположим теперь, что сильный ИИ в описанном выше смысле действительно будет когда-нибудь создан. Тогда, в принципе, люди окажутся более ненужными для дальнейшего саморазвития такого ИИ. В самом деле, для чего они, если ИИ по своим интеллектуальным возможностям превосходит все, что доступно людям? Более того, сильный ИИ может начать саморазвитие со столь высокой скоростью, что люди не только окажутся лишними в этом процессе, но и принципиально не смогут за ним уследить и понять происходящее. Будущее для людей становится полностью непонятным и непредсказуемым. Эта ситуация и называется технологической сингулярностью [1]. Термин

был введен Вернором Винджем в 1993 г. [6], хотя похожие идеи высказывались неоднократно и раньше (см. для обзора [1, 6]). Ожидание технологической сингулярности порождает тревожные настроения, получившие отражение в различных публикациях, как, например, в статье Билла Джоя с характерным названием «Why the future doesn't need us» [7].

Наиболее известна демографическая сингулярность, которая еще в 60-х годах прошлого века была обнаружена рядом авторов: Х. фон Форстером и др. [8], И.С. Шкловским [9] и другими. Кривая роста народонаселения Земли до примерно 1970-го года оказывается приблизительно гиперболой, уходящей в бесконечность между 2025 и 2030 годом. Другие эволюционные кривые с обострением, приводящие к сингулярностям, тоже обычно описываются гиперболами с различными показателями. В отличие от таких эволюционных сингулярностей, технологическая сингулярность прямо ни с какими бесконечностями не связана. Термин «сингулярность» в последнем случае является метафорой, и означает скорее весьма критическую ситуацию, в которой может оказаться человечество, если сильный ИИ реально когда-нибудь появится.

Основной вопрос, который возникает в связи с концепцией ТС, состоит в том, когда это можно ожидать, и из каких соображений надо определять эту дату. В книге [1] дается следующий ответ на этот вопрос, который разделяют и другие сторонники ТС: Как только мощность коммерческих компьютеров, выраженная в операциях в секунду (в оригинале - количество операций в секунду за одну тысячу долларов), превзойдет совокупную вычислительную мощность мозга всего человечества, сильный ИИ будет создан, и технологическая сингулярность станет реальностью. Нетрудно видеть, что эта идея адресует вычислительную мощность не отдельно взятого человеческого мозга, а именно всего человечества, что коррелирует с определением сильного ИИ, которого мы придерживаемся в статье.

Для прогноза мощности компьютеров Рэй Курцвейл [1] использует так называемый закон Мура, в соответствии с которым вычислительная мощность компьютеров удваивается каждые полтора—два года. Вычислительную мощность мозга Курцвейл оценивает следующим несложным образом. Количество нейронов мозга, порядка 10^{11}, умножается на количество синаптических связей одного нейрона, масштаба тысячи, и на частоту срабатывания одной синаптической связи, около сотни Гц. Получается порядка 10^{16} операций в секунду на один мозг. Если население Земли оценить как десять миллиардов, то результирующая вычислительная мощность всего человечества будет порядка 10^{26}. Эта величина и сравнивается с кривой Мура. Точка пересечения падает приблизительно на 2045 год - это и есть прогноз даты технологической сингулярности от Рэя Курцвейла.

Насколько обоснованным является такой прогноз? Не вполне очевидно, что закон Мура сохранит свою силу в течение достаточно длительного времени, хотя пока новые точки хорошо ложатся на кривые, использованные Курцвейлом в 2005 году, когда писалась книга [1]. Но вовсе не закон Мура, с нашей точки зрения, является самым слабым местом в предсказании ТС. За этим прогнозом стоят, как минимум, еще три плохо обоснованных предположения и одно полностью не понятое обстоятельство. Плохо обоснованные предположения суть следующие: 1.Возможность создания сильного ИИ определяется наличием компьютеров достаточной мощности; 2.Вычислительная мощность мозга определяется совокупным быстродействием синаптических связей нейронной сети; 3.Вычислительная мощность мозга вообще может оцениваться на основе аналогии `мозг - это классический компьютер'.

Полностью не понятым обстоятельством является теорема Пенроуза об искусственном интеллекте, которая вообще запрещает реализацию всех без исключения ментальных способностей человека на базе архитектуры клас-

сического компьютера. В последующих разделах статьи мы подробно обсудим все эти слабые места в аргументации сторонников сильного ИИ и ТС, причем особое внимание уделим разъяснению смысла теоремы Пенроуза, после чего критически рассмотрим интерпретацию следствий теоремы Пенроуза, предложенную самим Роджером Пенроузом.

Недооценка фактора программного обеспечения в создании сильного ИИ

«За прошедшие 15 лет "разум" наших электронных вычислительных машин улучшился в миллион раз... В течение нескольких следующих десятилетий следует ожидать увеличения характеристик "разума "машин еще по крайней мере в несколько десятков тысяч раз. "Разум" таких машин по основным параметрам будет *заведомо* превосходить разум человека». Это было сказано Иосифом Самуиловичем Шкловским в его знаменитой книге «Вселенная, жизнь, разум» издания 1975 года. С тех пор прошло более 37 лет, и мощность компьютеров за это время возросла вовсе не в несколько десятков тысяч раз, о чем писал Шкловский, а более чем в миллиард (!) раз (от машин серии БЭСМ-6 и Эльбрус и их зарубежных аналогов с скоростью до 10^7 флоп, до самой мощной современной супер-ЭВМ Titan, имеющей скорость $1.8 \cdot 10^{16}$ флоп).

Чтобы создать сильный ИИ, мало иметь достаточно мощное компьютерное железо. Надо знать, как это сделать. Нужны соответствующие методы, нужно программное обеспечение, нужно понимание того, какую именно задачу надо решить для создания сильного ИИ. Но программное обеспечение гораздо более консервативно, чем аппаратное обеспечение.

Синтетическое направление в разработке ИИ

Состояние синтетического направления характеризуют несколько ярких примеров. Программы для проведения аналитических вычислений (системы компьютерной

алгебры) являются типичным примером современных систем искусственного интеллекта. Одной из лучших таких современных систем является программа Maxima. При этом первые версии программы Maxima (тогда она называлась Macsyma) были разработаны еще в 1972 году, и основное вычислительное ядро системы с тех пор практически не менялось.

Windows-версия популярного текстового процессора Microsoft Word появилась в 1989 году, 23 года назад. Компьютеры за это время стали почти в миллион раз быстрее, но функциональность Word практически не изменилась.

В популярных изданиях или в сети нередко можно встретить заголовки вроде «Робот-учёный делает открытия без помощи человека» (см. например [10]). Творческие способности аккумулирует в себе программист, который ставит задачу и определяет, как ее надо решать.

Пожалуй, действительный прогресс наблюдается в направлении, которое известно как «искусственная жизнь - адаптивное поведение (аниматы)» (см. например [11]). Однако и здесь результаты пока довольно ограниченные.

Нетрудно видеть, что стремительно возрастающее совершенство компьютерной техники, выражаемое кривой Мура, расходуется пока почти исключительно на развитие пользовательского интерфейса, миниатюризацию и телекоммуникации. Создание сильного ИИ такое развитие вовсе не приближает. Еще одной бедой является почти катастрофический рост объема (в байтах, в строках программного кода) программного обеспечения без радикального роста его интеллектуальности. Это говорит о том, что даже небольшое продвижение в интеллектуальности оплачивается экспоненциальным ростом объемов кода программ.

Обратная инженерия мозга

Хотя в направлении обратной инженерии мозга ведется реальная работа, в частности в России под руково-

дством В.Л. Дунина-Барковского в лаборатории обратного конструирования мозга имени Дэвида Марра [12], задача кажется очень сложной и перспективы этого направления не вполне ясны. Пессимизм вселяют результаты работ по моделированию нервной системы нематоды Caenorhabditis elegans (сокращенно C. elegans) [13].

Нематода C. elegans - крохотный червячок с длиной тела всего около миллиметра, причем C. elegans представлен особями трех полов: мужскими, женскими и гермафродитами. Мужские и женские особи имеют нервную систему, состоящую примерно из тысячи нейронов, но нервная система гермафродитов состоит всего точно из 302 нейронов. При всей простоте нервной системы, нематода демонстрирует сложный репертуар поведений: навигация, поиск пищи, спаривание, обучение, социальное поведение, сон. Гермафродитные особи C. elegans являются очень удобным объектом для изучения и моделирования работы нервной системы. Вся нервная система C. elegans полностью и точно картирована, каждый нейрон имеет свое имя, причем для моделирования нервной системы, состоящей всего из 302 нейронов, проблема мощности компьютера заведомо не играет никакой роли. Однако, несмотря на то, что работы по моделированию нервной системы C. elegans ведутся с начала 1990-х годов, результаты крайне ограничены. Ничего похожего на полноценную работающую модель нейронной системы C. elegans до сих пор получить не удалось, имеются только некоторые ограниченные результаты в моделировании управления движением тела нематоды.

Интересно отметить, какие именно характерные трудности встали на пути решения задачи. Во-первых, оказалось, что даже самого полного картирования нейронной системы мало. Нужно знать силу синаптических связей (пороги возбуждения каждого нейрона через каждую синаптическую связь), но они остаются неизвестными, и измерить их пока не удается. Во-вторых, мало моделировать нервную систему. Чтобы понять, насколько полноценно и

адекватно работает модель, нужно либо создать полноценного робота C. elegans, либо полную компьютерную модель тела (и вообще всего организма) C. elegans, чтобы наблюдать результаты работы нервной системы. В работе [13] и некоторых других статьях по этой тематике исследователи пошли по пути создания компьютерной модели тела. Но, однако, и этого оказалось мало. Тело должно существовать в среде обитания, и ее тоже нужно моделировать. В [13] с помощью уравнения Навье-Стокса моделировались жидкие среды разной вязкости.

Мозг человека содержит порядка ста миллиардов нейронов вместо 302 нейронов C. elegans, поэтому задача обратного конструирования мозга должна быть на много порядков сложнее. При этом имеется также задача сопряжения компьютерной модели мозга либо с телом робота-андроида, представляющего полноценную модель тела человека, либо с компьютерной моделью тела, но тогда и с компьютерной моделью всей среды обитания, которой для человека является вся Вселенная. Последний вариант показывает, что моделирование процессов одного мозга неожиданно оказывается эквивалентным моделированию всей Вселенной, что вряд ли возможно. В варианте сопряжения модели мозга с телом робота-андроида придется еще решить вопрос о мотивах поведения и вообще существования для такого моделированного мозга. Мотивами поведения нормального человека является его чувственно-эмоциональная сфера и ощущение себя частью социума. Если перенести полностью такие мотивации на робота, то результатом будет просто искусственный человек, который вынужден будет и жить в реальном времени, в контакте с реальными людьми. Как удастся такому искусственному человеку примириться с наличием своих сверхспособностей?

Таким образом, как опыт развития синтетического направления конструирования ИИ, так и опыт обратной инженерии мозга показывает, что проблема состоит не

только, и даже не столько в том, что для создания сильного ИИ не хватает вычислительных ресурсов, сколько в том, что непонятно, как решать задачу. И даже, собственно, непонятно, какую задачу надо решать, поскольку мы не знаем, что такое чисто человеческая способность к пониманию, что такое свобода воли - то есть непонятно, что именно нужно перенести в компьютер.

Цитоэтология и быстродействие мозга

Как уже отмечалось, прогнозы Рэя Курцвейла в отношении даты наступления ТС основаны на оценке быстродействия мозга просто как суммарной максимальной скорости срабатывания всех синаптических связей мозга. Убеждение, согласно которому работа мозга как нейронной сети в основном и исчерпывает все основные функции, которые реализует мозг, называется нейронной парадигмой. С точки зрения этой парадигмы, для того, чтобы смоделировать работу мозга, достаточно смоделировать работу нейронной сети мозга. Однако мозг реализует еще один (как минимум) вид активности, который отвечает не за работу нейронной сети мозга, а за модификацию структуры этой нейронной сети. Наиболее очевидными процессами этого типа являются возникновение и исчезновение синаптических связей.

Примеров такого сложного поведения известно довольно много, и один из недавно обсуждавшихся в литературе относится к слизевику - одноклеточному колониальному существу [14]. Слизевик демонстрирует разные виды «интеллектуального поведения», среди которых, в частности, можно отметить способность проходить лабиринты, оптимизировать геометрическую форму колонии для достижения определенных целей и др. Но, пожалуй, самой удивительной оказывается способность к обучению (выработке условного рефлекса). Слизевики способны медленно перемещаться (подобно амебам), и было обнаружено, что влажный воздух заставляет их двигаться быстрее, а сухой -

наоборот, замедляет перемещение. Чередуя поток влажного и сухого воздуха с определенным периодом, была обнаружена интересная особенность: перед очередной подачей сухого воздуха слизевики снижали скорость. При периодической смене влажного и сухого воздуха колония запоминала последовательность этих перемен, и продолжала помнить ее несколько периодов, даже если смена потоков прекращалась. Считается, что мозг для запоминания информации использует изменение силы синаптических связей нейронной сети. Но чем запоминает слизевик, если у него вовсе нет нервной системы? Все сложное поведение слизевика, особенно описанный опыт с обучением, показывают, что существуют сложные внутриклеточные механизмы обработки информации, включая возможность обучения на основе использования внутриклеточной памяти. Этот круг явлений настолько богат и своеобразен, настолько отличается от того, что изучается стандартно понимаемой цитологией и любой другой наукой, имеющей отношение к биологии, биохимии или биофизике клетки, что его должна изучать новая, практически еще не оформившаяся наука, которую можно назвать *цитоэтологией*. Термин был введен В.Я. Александровым в статье [15].

Биологическая эволюция устроена таким образом, что раз обретенные находки и решения не теряются, но оказываются встроенными в весь последующий эволюционный процесс. Генетический код, появившийся у бактерий, без всяких изменений используют и высшие животные; многоклеточные живые организмы есть ни что иное, как сложно организованные колонии узкоспециализированных одноклеточных организмов - клеток т. д. Это свойство эволюционного процесса связано с такими понятиями, как аддитивность и консерватизм эволюции [16, с. 27]. Поэтому следует ожидать, что механизмы внутриклеточного управления, соответствующие уровню цитоэтологии, появившись в одноклеточном мире, унаследованы и клетками высших многоклеточных организмов. Индивидуальные

нейроны должны проявлять сложные формы поведения, не сводящиеся только лишь к функции нейрона как порогового переключателя в нейронной сети.

Действительно, такие виды активности, оказывается, уже давно были открыты. В статье В.Г. Режабека [17] в эксперименте над нейрорецептором растяжения речного рака было убедительно показано, что одиночный изолированный нейрон способен к обучению. Примечательно не только то, что одиночный нейрон способен обучаться, но ситуация, в которой происходило обучение нейрона в опытах В.Г. Режабека [17], имеет очень мало общего с реальными жизненными ситуациями, с которыми приходится сталкиваться нейрону. Поэтому, помимо способности к обучению, в этих опытах была продемонстрирована еще и удивительная гибкость поведения нейрона. Возникает естественный вопрос: какая же эффективная скорость вычислений, выраженная в операциях в секунду, может отвечать этим внутринейронным процессам управления? Чтобы ответить на этот вопрос нужно, естественно, понимать, где на субклеточном уровне может помещаться механизм «внутриклеточного сознания».

Остановимся на одной из наиболее широко обсуждаемых возможностей - системе так называемых микротрубочек, из которых построена значительная часть цитоскелета клетки. Мы будем опираться, в основном, на книги Роджера Пенроуза [2, 3], и его же недавнюю статью [18], написанную совместно с Стюартом Хамерофф'ом.

Цитоскелет клетки представляет собой систему тонких белковых волокон, пронизывающих всю клетку. Цитоскелет - динамичная, изменяющаяся структура, в функции которой входит поддержание и адаптация формы клетки к внешним воздействиям, экзо- и эндоцитоз, обеспечение движения клетки как целого, активный внутриклеточный транспорт и клеточное деление [19]. Цитоскелет сложен структурами нескольких типов, для нас основной интерес представляют так называемые микротрубочки. Часть цито-

скелета состоит из пучков таких микротрубочек. Каждая микротрубочка представляет собой, действительно, полую трубку с внешним диаметром около 25 нм, стенки ее сложены ровно из 13 рядов молекул белка-тубулина, причем молекулы в стенке уложены в правильную кристаллическую структуру. Каждая молекула тубулина представляет собой димер, состоящий из двух частей, называемых α-тубулином и β-тубулином. Молекула тубулина может находиться в двух конформациях, различающихся расстоянием между α- и β-тубулином и имеющих разный дипольный электрический момент. По этой причине молекулы тубулина могут играть роль битов с двумя состояниями. Более того, соседние молекулы тубулина в микротрубочке взаимодействуют между собой своими электрическими моментами, и благодаря правильной кристаллической структуре всей системы микротрубочка чрезвычайно напоминает клеточный автомат. Клеточный автомат, в свою очередь, может быть универсальным средством вычисления и управления (имеется доказательства возможности эмуляции универсальной машины Тьюринга клеточным автоматом). Следовательно, микротрубочки вполне могут оказаться тем универсальным вычислительным устройством, которое отвечает за сложное поведение клеток или автономных одноклеточных существ [20].

В статье [18] на основе обзора нескольких исследований утверждается, что микротрубочки демонстрируют набор резонансных частот, которые можно наблюдать в спектрах поглощения и излучения электромагнитных волн, в диапазоне от нескольких кГц то примерно 10^7 Гц. При этом именно наивысшая частота 10^7 Гц больше всего похожа на основную частоту колебаний дипольного момента молекулы тубулина, входящей в состав кристаллической решетки микротрубочки. В этом случае 10^7 Гц и есть скорость срабатывания одной ячейки клеточного автомата, а быстродействие всего автомата можно получить умножив эту частоту на количество ячеек в нем. Поскольку один

нейрон в составе микротрубочек содержит порядка 10^8 молекул тубулина, то полное быстродействие одного нейрона в этих терминах оказывается 10^{15} операций в секунду, а быстродействие всего мозга с его сотней миллиардов нейронов оказывается 10^{26} операций в секунду. Это на десять порядков превышает оценку быстродействия мозга, предлагаемую Рэем Курцвейлом для прогноза даты ТС, и вызывает очень большие сомнения, что закон Мура будет исправно действовать вплоть до столь огромных величин.

Помимо микротрубочек в клетке имеются и другие кандидаты на локализацию информационных процессов. Поэтому основная оценка быстродействия мозга, как 10^{16} операций в секунду, представляется обоснованной очень слабо.

Аналогия мозг-компьютер и квантовые моды работы мозга

Методика прогноза даты ТС недвусмысленно подразумевает, что быстродействие компьютеров и мозга вообще можно сравнивать, то есть что мозг можно рассматривать как обычный классический компьютер, быстродействие которого может быть выражено в количестве операций в секунду. Между тем, если на некотором уровне (пусть даже весьма глубоком) в мозге происходит обработка информации в квантовом режиме, подобно тому, как это делается в квантовом компьютере, то такое сопоставление полностью утрачивает смысл.

Основное возражение против такой возможности со стороны Рэя Курцвейла состоит в том, что мозг — это место, которое очень плохо подходит для существования так называемой квантовой когерентности, необходимой для проведения квантовых вычислений. Он цитирует слова Сета Ллойда (Seth Lloyd), известного специалиста по квантовой теории и квантовой информатике, из его интервью электронному журналу Nano Magazine: «Мозг - это теплое

и влажное место. Это очень неудачное окружение для использования квантовой когерентности».

Действительно, большинство существующих прототипов квантовых компьютеров требуют для своей работы экстремально низкотемпературных условий и высокой степени изоляции устройства от окружения, и эту особенность конструкции Курцвейл неявно переносит на все вообще возможные устройства этого класса. Но такая экстраполяция вовсе не является оправданной и обоснованной. Большинство современных прототипов квантовых компьютеров используют либо манипуляцию спиновыми состояниями (спинтроника), либо сверхпроводящие элементы. Спиновые состояния, как правило, являются очень хрупкими (подвержены декогеренции), так как отделены от окружения лишь очень малой энергетической щелью (либо не отделены практически вовсе). Поэтому сохранение таких состояний, вообще говоря, требует очень низких температур и высокой изоляции от окружения. Сверхпроводящие элементы нуждаются в низкотемпературных условиях по той простой причине, что настоящая высокотемпературная сверхпроводимость (при комнатной температуре) пока неизвестна. Однако, элементы квантовых вычислительных устройств могут использовать совсем другие физические принципы, которые позволят обходиться без криогеники для сохранения квантовой когерентности. Более того, недавно уже было продемонстрировано двух-кубитное квантовое вычислительное устройство, хоть и являющееся спинтронным, но работающее при комнатной температуре [21]. В этом случае для с борьбы с декогеренцией была использована (и впервые продемонстрирована на практике) активная квантовая коррекция квантовых операций, но и без такой активной квантовой коррекции устойчивое существование квантово-перепутанных состояний при нормальных условиях вовсе не является чем-то необычным. Приведем несколько примеров.

Тривиальный пример представляет атом гелия.

Нижняя оболочка атома заселена двумя электронами, но состояние каждого из этих электронов не описывается какой-то определенной волновой функцией. Волновую функцию имеют только оба электрона вместе, и такая волновая функция является перепутанным квантовым состоянием двух электронов. Это перепутанное состояние без всяких проблем существует не только при комнатной температуре, но и приблизительно до двадцати пяти тысяч градусов, когда электроны начинают переходить в возбужденное состояние и может произойти ионизация атома. Менее тривиальный пример представляют молекулы красителей. Этот пример был использован Ричардом Фейнманом для иллюстрации фундаментальных положений квантовой механики в его знаменитом курсе лекций [22, с. 195-196].

Таким образом, в существовании квантовой спутанности даже для макромолекул при нормальных условиях, в растворе или в составе твердого тела, нет ничего необычного и, тем более, невозможного. Эволюция вполне могла бы найти способ использования таких состояний внутри живой клетки, если бы только это давало какие-то селективные преимущества.

Можно ли указать место в нейроне, где могли бы существовать и как-то использоваться такие квантово-перепутанные состояния? Все те же упомянутые выше микротрубочки являются одним из вполне разумных кандидатов на эту роль [2, 3, 18]. Микротрубочки не просто погружены в цитоплазму клетки, но стенка микротрубочки внутри и снаружи покрыта пленкой воды, причем такой воды, которая находится в состоянии, близком к кристаллическому. Эта пленка кристаллической воды очень хорошо изолирует стенку микротрубочки, сложенную молекулами тубулина, от внешнего окружения. Поэтому не видно ничего невозможного в том, что молекулы тубулина существуют не просто в одной из двух своих возможных конформаций, но в когерентной суперпозиции обеих конформаций, при этом соседние тубулины могут образовывать

еще и перепутанные состояния. Тогда клеточный автомат микротрубочки будет квантовым клеточным автоматом, который может реализовать квантовый уровень управления нейрона. Есть и другие места в клетке, которые могут служить кандидатами на локализацию квантово-информационных процессов.

Таким образом, существование квантовых процессов обработки информации в мозге вовсе не исключено, хотя и не доказано прямыми наблюдениями. Однако, как мы увидим ниже, непрямое доказательство большой роли квантово-информационных процессов в работе мозга существует, и связано с теоремой Пенроуза об ИИ. Если, действительно, мозг в ходе своей работы реализует что-то подобное квантовым вычислениям, то для того, чтобы превзойти мозг просто по скорости обработки информации, нужен не компьютер в обычном понимании, а некоторое устройство, в состав которого входят модули квантовых вычислений. После почти тридцати лет развития идеи квантового компьютинга, создано первое полнофункциональное квантовое вычислительное устройство, поддерживающее алгоритмы квантовой коррекции кода, которые позволяют бороться с декогеренцией. Хотя устройство это состоит всего из двух кубитов [21], может быть эти первые два кубита положат начало кривой Мура в отношении квантовых компьютеров.

Нейросетевая парадигма работы мозга совсем не обязательно является истиной в последней инстанции. Противоположная крайность состоит в том, что нейросетевая активность не есть носитель сознания, но есть лишь инструмент сознания, интерфейс, связывающий сознание с окружающей действительностью. «Само» сознание (или его жизненно важные компоненты) живет на некоторых более глубоких (субнейронных, квантовых?) уровнях организации мозга. Возможно, истина лежит где-то посередине.

Наконец заметим, что возможное наличие субнейронных уровней обработки информации очевидным обра-

зом резко усложняет задачу обратного конструирования мозга, в особенности, если на этих уровнях присутствует квантовая обработка информации.

«No-go» теорема Роджера Пенроуза об искусственном интеллекте

Закон сохранения энергии (первое начало термодинамики) запрещает создание вечного двигателя первого рода. Второе начало термодинамики запрещает создание вечного двигателя второго рода. Очень похоже, что роль, аналогичную первому и второму началам термодинамики в отношении вечных двигателей, относительно возможностей ИИ играет теорема Пенроуза об искусственном интеллекте.

Содержание теоремы сводится к утверждению, что какой бы мощностью ни обладало устройство, имеющее архитектуру конечного автомата (компьютера в современном понимании), человеческое мышление имеет некоторые возможности, недоступные такому устройству. Ни один компьютер не может превзойти мышление человека во всех отношениях независимо от его мощности, так как теорема говорит о том, что в некотором отношении человеческое мышление обязательно будет сильнее. Сосредоточимся на смысле теоремы Пенроуза, которая осталось совершенно непонятой сторонниками сильного ИИ и ТС.

Начать нужно с первой теоремы Гёделя о неполноте (см. например [23, с. 188]), на которую часто ссылаются просто как на теорему Гёделя. Смысл теоремы состоит в следующем. Пусть имеется любая непротиворечивая аксиоматическая система, содержащая в себе формальную арифметику. Тогда в этой системе существует осмысленное утверждение, которое нельзя ни доказать, ни опровергнуть средствами этой системы. Более того, доказательство теоремы имеет конструктивный характер в том смысле, что это утверждение строится в явном виде, и является истинным по построению (на метаматематическом уровне, см.

[23, с. 188]). Для доказательства теоремы используются две фундаментальные идеи: так называемая Гёделевская нумерация и диагональный метод Кантора. Сама теорема Гёделя в высшей степени нетривиальна, содержит в себе множество тонкостей (некоторые детали обсуждаются в нашей статье [24]), но к настоящему времени чрезвычайно подробно исследована и не вызывает никаких сомнений.

Структура любого конечного автомата (компьютера) может быть описана конечным образом просто в силу конечности этого устройства. Это описание аналогично конечному набору аксиом некоторой формальной системы. Как и следует ожидать, в полной аналогии с теоремой Гёделя, для такой системы (машины) можно явно построить некоторое истинное утверждение, истинность которого не может быть доказана (точнее говоря, вычислена) данным конечным автоматом. В этом заключается смысл теоремы Гёделя-Тьюринга для конечных автоматов.

Теорема Гёделя-Тьюринга ничего не говорит об интеллектуальных возможностях человека. Именно это обстоятельство не понято сторонниками сильного ИИ. Здесь необходим еще один весьма нетривиальный шаг, и Роджер Пенроуз этот шаг делает. Рассуждение Пенроуза представляет собой доказательство от противного. Предположим, что создан суперкомпьютер, имеющий архитектуру конечного автомата, который реализует, как минимум, все методы математических рассуждений, которыми владеет человечество (сильный математический ИИ). Однако, в силу теоремы Гёделя-Тьюринга, для данного суперкомпьютера, как и для любого конечного автомата, любой человек, понимающий теорему Гёделя-Тьюринга, *используя математическое рассуждение*, зафиксированное в этой теореме, может явно построить утверждение, про которое ему будет точно известно, что оно истинно (по построению), хотя для данного суперкомпьютера его истинность недоступна. Мы получили противоречие: предположив, что суперкомпьютер владеет всеми методами математических рассуждений,

которыми владеют люди, мы немедленно указали математическое рассуждение, которое доступно любому человеку, понимающему теорему Гёделя-Тьюринга, но недоступное этому компьютеру. То есть мы доказали, что компьютер владеет не всеми методами математических рассуждений людей. Противоречие доказывает, что исходное предположение было неверным, следовательно суперкомпьютер, владеющий всеми математическими способностями людей, невозможен, тем самым некоторые способности людей остаются за пределами достижимости любого вычислительного устройства - конечного автомата. *Сильный ИИ невозможен ни для каких компьютеров на основе архитектуры конечного автомата.*

Из теоремы Гёделя-Тьюринга следует так же, что мозг человека, способного понять теорему Гёделя-Тьюринга (назовем такого человека *математиком*), сам не является конечным автоматом. Действительно, предположим, что мозг математика - это некоторый конечный автомат. Тогда, используя теорему Гёделя-Тьюринга, математик может построить истинное утверждение, истинность которого он не может установить с использованием его собственного мозга, как конечного автомата. Тем самым, такой математик понимает то, что он понять не может. Противоречие доказывает, что исходное предположение было неверным, и мозг математика не является конечным автоматом. Следовательно, аналогия «мозг - это компьютер», неверна для мозга любого отдельно взятого математика. Это утверждение можно рассматривать как альтернативную формулировку теоремы Пенроуза.

Интересно, что этим способом невозможно доказать, что мозг любого наперед заданного человека не является конечным автоматом. Теорема справедлива только либо в отношении математической способности всех людей, взятых вместе (среди которых есть, разумеется, и математики), либо в отношении отдельных личностей, понимающих или способных понять теорему Гёделя-Тьюринга.

Но дальше логика Роджера Пенроуза такова. Математические способности людей представляют только частный случай способностей мозга, отличающийся от других способностей тем, что здесь анализ соотношения способностей мозга и компьютера удается абсолютно строго довести до конца. Поскольку в отношении математических способностей точно доказано превосходство человека над машиной, то, по аналогии, человек, может обладать и другими способностями, недоступными конечному автомату. Просто анализ в других случаях довести до конца труднее, и, скорее всего, источник всех этих «невычислимых» способностей один. Действительно, многое указывает на существование таких способностей, и Роджер Пенроуз приводит множество примеров. Качественный анализ этого типа проводился и задолго до книг Пенроуза, например в книге Хьюберта Дрейфуса [25] (конец шестидесятых годов XX века).

Таким образом, человек - не компьютер не только в смысле его математических способностей (что, фактически, доказано), но и во многих других отношениях. Как написали братья Стругацкие в повести «Беспокойство»: «все фундаментальные идеи выдумываются..., они не висят на концах логических цепочек». Теорема Пенроуза указывает на то, что эта универсальная способность людей выдумывать и не может быть формализована на основе архитектуры конечного автомата. Автоматы ходят только вдоль логических цепочек, поэтому новые фундаментальные идеи, которые не висят на их концах, им недостижимы. Но где они висят, эти фундаментальные идеи, как люди до них добираются? Мы этого не знаем.

Надо отметить, что анализ книги Рэя Курцвейла о технологической сингулярности [1] не идет дальше теоремы Гёделя-Тьюринга. Имя Пенроуза в контексте теоремы Гёделя-Тьюринга здесь даже не упоминается, и это при том, что с книгами Роджера Пенроуза Рэй Курцвейл точно знаком (что видно по обильному цитированию по другим

поводам в других местах книги [1]). Это означает, что теотеорема Пенроуза просто не была понята. Например, в разделе «The Criticism from the Church-Turing Thesis» в книге [1] можно встретить такое рассуждение (не дословно): Компьютер - конечный автомат, поэтому для него существуют Гёделевские утверждения. Это несомненно. Но и мозг - тоже конечный автомат (это для Рэя Курцвела ясно), поэтому и для него тоже существуют Гёделевские утверждения. В этом смысле, компьютеры ничуть не хуже мозга, поэтому нет ничего удивительного в том, что компьютер может полностью воспроизвести работу мозга.

Следующий шаг от теоремы Гёделя-Тьюринга, который приводит к выводу о том, что мозг не может быть конечным автоматом, Рэй Курцвейл не сделал сам и не увидел его у Роджера Пенроуза. Напротив, в книге Хьюберта Дрейфуса [25] представлено ясное понимание того, что мозг не может быть просто компьютером. Собственно, в этом состоит основной смысл книги [25]. Интересен ответ на вопрос о том, почему мозг может не быть компьютером. Хьюберт Дрейфус формулирует его примерно так: машины обрабатывают информацию, а человек работает со смыслами. Вовсе не очевидно, что человеческие смыслы, чем бы они ни были на самом деле, могут быть закодированы информацией. Поэтому, возможно, мысль – не вычисление в обычном смысле.

Природа невычислительной активности мозга по Роджеру Пенроузу

Если мозг - не компьютер, то что это? В чем физически заключается корень невычислительной активности мозга? Прежде всего Пенроуз показывает, что архитектура мозга как нейронной сети не имеет отношения к делу. Этот аспект проблемы связан с различием «нисходящих» и «восходящих» способов программирования. Под нисходящими способами понимаются традиционные методы, основанные на процедурном программировании, когда сначала

пишется головная процедура программы, она обращается к подпрограммам, те обращаются к подпрограммам более низкого уровня и т.д., пока алгоритм не будет полностью детализирован вплоть до уровня машинных команд, не требующих дальнейшей детализации. Большинство программных систем до сих пор создаются именно этим способом. Восходящее программирование характерно для систем, в которых реализуется та или иная форма «обучения», и, на первый взгляд, какой-либо заранее определенный жесткий алгоритм, предназначенный для решения определенной группы задач, отсутствует. Наиболее типичным примером этого типа программирования являются искусственные нейронные сети. В любой самообучающейся искусственной системе (из числа известных в настоящее время) может быть найдено жесткое алгоритмическое ядро. Так что использование искусственных нейронных сетей какой угодно сложности и любых других методов восходящего программирования вовсе не выводит за пределы понятия конечного автомата, и нейросетевая структура мозга, взятая сама по себе, тоже не имеет отношения к его невычислительной активности.

Значит, причину невычислительной активности мозга надо искать в чем-то другом. Логически остаются две возможности: либо этим чем-то может быть особая физика, управляющая работой мозга, либо источник невычислительной активности лежит вне мозга и соответствующая способность мозга является следствием того, что мозг является открытой системой. Обе эти возможности рассматриваются Роджером Пенроузом).

На идею о том, что причина невычислительной активности мозга может быть следствием открытости мозга как системы и может лежать за пределами самого мозга, Пенроуз возражает двумя способами. Во-первых, если только предполагать, что любой конечный фрагмент окружения человека описывается вычислимой физикой, то он, формально говоря, может быть отображен на работу ко-

нечного автомата, поэтому ничего невычислимого в нем не найдется. В этом смысле невычислимости вне мозга просто неоткуда взяться. Никакие контакты мозга с окружением к невычислимости тогда не приведут. Во-вторых, если мы предполагаем, что вне человека всё-таки есть что-то, имеющее невычислимый характер, то мы, очевидно, соглашаемся, что невычислимые процессы в природе существуют, хотя бы в принципе. Но тогда нелепо предполагать, что они локализуются в каком-то месте, отличном от самого мозга, так как именно мозг является самой сложной из известных нам природных систем. То есть источник невычислительной активности мозга нужно искать только внутри мозга, никак не снаружи. Далее Пенроуз детально рассматривает разные типы физических процессов в качестве потенциальной основы невычислительной активности внутри мозга (этому посвящена вторая половина каждой из книг [2, 3]). Пенроуз приводит несколько примеров, из которых следует, что в такой физике, в принципе, нет ничего невозможного. Пенроуз тщательно пересматривает в поисках невычислимости всю известную физику. Сначала он исключает классическую механику и классическую теорию поля, включая теорию гравитации (общую теорию относительности ОТО). Затем он рассматривает классическую статистическую физику. Задача сводится к моделированию ансамблей. Это снова не приводит к каким-либо принципиальным проблемам, связанным с вычислимостью. Ансамбли можно моделировать с использованием программных генераторов случайных чисел, которые могут следовать чисто случайному поведению с практически сколь угодно высокой точностью. Таким образом, вся классическая физика, как обыкновенная, так и статистическая, исключается полностью.

На этом этапе возникает, ни много ни мало, непрямое доказательство того, что мозг в процессе мышления существенным образом использует квантовые процессы. Действительно, если вся классическая физика, как источ-

ник невычислительной активности мозга, исключена, то остается искать такую активность в квантовой физике (как минимум). В случае квантовой активности мозга тоже требуется ее прямое обнаружение, одного только доказательства через теорему Пенроуза не достаточно.

После классической физики Роджер Пенроуз рассматривает квантовую физику, и снова не обнаруживает здесь ничего невычислимого. Даже если мозг - не классический, а квантовый компьютер в современном понимании этого термина, это никак не может объяснить его невычислительную активность. Мозгу мало быть просто квантовым компьютером, нужно нечто большее. В основе работы мозга должна лежать еще не известная невычислимая физика, ничего другого не остается! Таков вывод Роджера Пенроуза.

Окончательный вывод Роджера Пенроуза о необходимости новой физики для понимания работы мозга кажется более чем поразительным. Получается, что в новую физику (квантовую гравитацию?) может вести вовсе не только создание гигантских коллайдеров для исследования физики микрочастиц, или лишь ненамного менее гигантских и сложных телескопов для исследований в области космологии ранней Вселенной, гамма-барстеров и др., но наука о сознании. Мозг может быть таким же пробным камнем в поисках фундаментальной физики, как упомянутые ускорители и телескопы. Надо, конечно же, отметить, что Пенроуз среди ученых вовсе не одинок в предположении, что наука о сознании, так или иначе, должна занять важное место в фундаментальной объединяющей физической теории. Хорошо известны исследования М.Б. Менского в этом направлении [26, 27, 28], недвусмысленно в этом же духе высказался Андрей Линде в недавнем интервью по случаю получения им премии Fundamental Physics Prize Юрия Мильнера [29]. Хотя, быть может, вывод Пенроуза и не столь уж удивителен. Действительно, физика элементарных частиц адресует экстремально мелкие пространствен-

ные масштабы, космология - Вселенную в целом, то есть экстремально большие масштабы, а мозг представляет самую сложную известную структуру - то есть экстремальную сложность.

Невычислительная активность Вселенной в целом

Первое уточнение имеет довольно тривиальный характер и относится к утверждению Роджера Пенроуза, что вне мозга не надо искать невычислительную активность. Невычислительная активность, существующая вне мозга, и не связанная прямо с чьей-либо еще мыслительной активностью, очевидным образом существует. Это активность Вселенной в целом. Активность Вселенной в целом принципиально не может быть «вычислена» конечным автоматом, так как для этого потребовался бы автомат (компьютер) заведомо превосходящий Вселенную в размере, что невозможно, так как он сам должен быть частью Вселенной. Активность Вселенной в целом невычислима.

Квантовая реальность: вычислимость и симулируемость

Вычислимость квантовой теории - «наивный» подход. На первый взгляд, вывод о том, что вся квантовая физика является вычислимой, тоже не вызывает никаких сомнений. Действительно, состояния квантовых систем представляются векторами в гильбертовом пространстве. Эволюция систем в квантовой теории представляется унитарными преобра-зованиями состояний систем - векторов гильбертова пространства. Наконец, третьим и последним компонентом квантовой теории являются измерения. Измерения характеризуются вероятностями получения на выходе измерительной процедуры тех или иных значений наблюдаемых величин, а сами вероятности определяются проекционным постулатом Борна - фон Неймана. Таким образом, все вычисления, необходимые для предсказания поведения квантовых систем, формально

говоря, можно выполнить на обычном классическом компьютере. Например, по этой причине можно поведение квантового компьютера полностью симулировать классическим конечным автоматом. Более того, такие программы-симуляторы существуют [30]. Философски говоря, квантовая реальность допускает исчер-пывающее представление в классических вычислительных системах - компьютерах. Более точно, можно говорить о принципиальной возможности изоморфизма фрагментов квантовой реальности и компьютерных (вычислительных) моделей этих фрагментов реальности.

Изоморфизм квантовой реальности и ее симуляции классическим конечным автоматом. Сформулированное выше представление о вычислимости квантовой теории приводит к некоторым не совсем тривиальным вопросам. Если квантовый процесс можно взаимно однозначно отобразить на функционирование классического конечного автомата, то, с точностью до изоморфизма, квантовый процесс – это просто определенный способ работы конечного автомата. Но куда же тогда девается вся таинственность квантового мира? Где квантовые парадоксы, о которых столько сказано? Конечный автомат сводит квантовое поведение к чему-то простому и классическому, в чем ничего таинственного нет. Программа для компьютера не может быть таинственной. Более того, не имеем ли мы противоречия с доказанными теоремами о невозможности существования классических скрытых параметров в квантовой механике? Действительно, любой классический компьютер представляет собой устройство, описываемое локальными параметрами, которые можно рассматривать и как скрытые параметры по отношению к моделируемой системе. Ячейки памяти - триггеры компьютера - представляют собой обычные классические системы, обладающие четкой локализацией в пространстве и времени. Состояния триггеров тоже полностью локализованы в пространстве-

времени и соответствуют представлению о локальных классических переменных. И, тем не менее, система, составленная из таких классических устройств, сама являющаяся классической и локальной, демонстрирует квантовое поведение. Как такое может быть? Что здесь не так? Для того, чтобы разобраться с этим вопросом, полезно иметь перед глазами пример простейшей нетривиальной системы, поведение которой, как это принято считать, прямо противоречит наличию классических скрытых параметров. Это ЭПР-пара скоррелированных частиц со спином 1/2. Для такой системы справедлива теорема Белла [31], которая говорит о том, что если в основе поведения системы лежат локальные скрытые параметры, то при измерении проекций спинов частиц должны выполняться определенные неравенства, которые называются неравенствами Белла. Реальные же квантовые ЭПР-пары нарушают эти неравенства, что прямо противоречит существованию локальных класси-ческих скрытых параметров.

Классический компьютер вычислимым образом предсказывает нарушение неравенства Белла для ЭПР-пары. Однако, в связи с проблемой ИИ представляет интерес вопрос, может ли искусственное устройство *симулировать* работу мозга. Поэтому нас будет интересовать не просто вычислимость результата квантовой теории, но можно ли компьютер заставить работать *как* квантовую систему, то есть, не предсказывать, а *симулировать* поведение квантовой системы.

Чтобы понять в деталях, как нарушение неравенства Белла может быть согласовано с поведением обычного классического компьютера, мы написали простую программу, которая должна симулировать поведение скоррелированной ЭПР-пары. Логически, программа представляет собой структуру клиент-сервер. Сервером является та часть программы, в которой хранится состояние ЭПР-пары, и в которой физически реализованы процедуры, связанные

с преобразованием состояния при проведении некоторого измерения над этим состоянием. Два клиента *A* и *B* представляют измерения спинов, проводимые над частицами *A* и *B*. Программа работает следующим образом. Сначала формируется начальное синглетное состояние ЭПР-пары, которое записывается в сервер, и сервер посылает сигнал готовности состояния в клиенты. Это действие симулирует достижение частицами измерительных устройств. Получив сигнал, каждое из устройств должно провести измерение над состоянием объединенной системы двух частиц, которое хранится в сервере в единственном экземпляре. Очевидно, клиент должен обратиться к серверу за состоянием. Для этого каждый из клиентов посылает в сервер сигнал - запрос на измерение состояния частицы. Этот сигнал содержит информацию о том, для какой частицы и вдоль какой оси проводится измерение спина. Сервер обрабатывает этот сигнал в соответствии с проекционным постулатом. Для текущего состояния системы, записанного в сервере, он, используя датчик случайных чисел, генерирует результат измерения, и, в соответствии с полученным результатом, проводит необходимое преобразование состояния системы. Затем сервер возвращает результат измерения в тот клиент, который выдал запрос на измерение. Это есть завершение процедуры измерения спина - измерительное устройство получает результат измерения. Для каждого приготовленного состояния ЭПР-пары каждый клиент-измеритель обращается к серверу независимо от другого. Процедура повторяется многократно, так накапливается статистика для вычисления корреляций результатов измерений. В отношении измерения корреляций, компьютерная модель ведет себя в точности как реальная ЭПР-пара. Неравенство Белла нарушено. Всё как в натурном эксперименте. Возникают два вопроса. Q1 Получили ли мы противоречие с теоремой Белла, симулировав поведение квантовой скоррелированной пары классическим локальным устройством? Q2 Получили ли мы на самом деле исчерпы-

вающую симуляцию ЭПР-пары?

На вопрос Q1 можно уверенно дать отрицательный ответ. Никаких противоречий здесь нет. Всё дело в том, как в точности понимает локальные классические скрытые параметры теорема Белла. В доказательстве теоремы Белла (см. [31]) центральным пунктом является предположение, что акты измерения могут быть разделены пространственно-подобным интервалом, следовательно, они не могут оказывать причинное влияние друг на друга. Пусть a и b - направления, вдоль которых проводится измерение спина частиц, соответственно A и B. Тогда из причинной независимости результатов измерений следует, что направление b не может оказать влияние на результат измерения A, и наоборот. В компьютерной модели роль скрытых переменных λ играет генератор случайных чисел. В нашей компьютерной модели это условие Белловской локальности определенно не выполняется. Таким образом, теорема Белла не запрещает существование классических скрытых параметров как таковых, но запрещает существование скрытых параметров, локализованных в областях, разделенных пространственно-подобными интервалами. Поэтому существование локальной классической компьютерной модели, нарушающей неравенства Белла, не противоречит теореме Белла.

И, кстати, здесь немедленно возникает новый вопрос. Раз пространственно-временные связи и соотношения вообще нерелевантны понятию конечного автомата, но, очевидным образом, они имеют самое прямое отношение к поведению реальных квантовых объектов вроде ЭПР-пар, разнесенных на пространственно-подобные интервалы, то, значит, представленная вычислительная реализация «квантовой реальности» упускает что-то очень существенное. Что именно? Действительно, в реальном пространстве-времени вычислительными средствами симулировать пространственно-временное поведение квантовых систем, вообще говоря, невозможно (при том, что нарушение нера-

венства Белла возможно). Представим себе, что система клиент-сервер реально разнесена в пространстве, и расстояние между сервером (размерами которого мы пренебрегаем) и клиентами, представляющими измерения, одинаково для обоих клиентов и равно L. В случае реальной ЭПР-пары измерение спинов происходит «мгновенно» (в пределе длительность может быть как угодно мала по сравнению с временами перемещения частиц). Напротив, в компьютерной симуляции время измерения принципиально конечно: клиент-измеритель должен отправить запрос к серверу-состоянию, сервер должен отправить ответ клиенту, и клиент должен ответ получить. Ответ на вопрос Q2 отрицательный: исчерпывающую симуляцию ЭПР-пары мы не получили. Пространственно-временные соотношения, характерные для реальной ЭПР-пары, в нашей модели нарушены. То есть, симулируемость поведения квантовых систем в реальном пространстве-времени не сводится к вычислимости квантовой теории.

Никакими силами невозможно классическую вычислительную систему заставить вести себя в реальном пространстве полностью как существенно распределенный в пространстве квантовый объект.

Мозг, если он действительно опирается в своей работе на квантовые процессы (что кажется неизбежным по теореме Пенроуза), вполне может являться распределенной квантовой системой, подобной ЭПР-паре (но в миллиарды раз сложнее). Поэтому для того, чтобы говорить о вычислимости квантовых процессов мозга, надо представлять себе реализацию модели мозга не в реальном, а в виртуальном пространстве-времени, подобно тому, как это было описано выше для ЭПР-пары. Но при этом модель мозга должна обрабатывать сигналы из реального пространства времени и «жить» в реальном пространстве-времени, чтобы иметь какую-то связь с действительностью. Как это возможно? Для такого мозга должно быть два пространства-времени - внутреннее, виртуальное, в котором модели-

руются квантовые корреляции, и внешнее, реальное, в котором мозг «живет». Не возникнет ли неустранимых противоречий в попытке совместить эти две вещи в одной модели? Например, какое пространство-время должны адресовать модели внутренних сенсоров мозга, сигнализирующих о головной боли (и должна ли у компьютерной модели мозга болеть голова, если она представляет мозг, страдающий головной болью?). Как классические подсистемы мозга (вроде внешнего поведения нейронной сети как системы классических пороговых переключателей) в модели должны взаимодействовать с виртуальным пространством, в котором описываются корреляции?

Неинформационная природа квантовых состояний. Как уже упоминалось, Хью-берт Дрейфус писал [25], что смыслы, которыми оперирует человеческое сознание, могут не иметь простой информационной природы.

Из теоремы Пенроуза и анализа вычислимости классической физики следует, что в классической (не квантовой) физике источник невычислимой активности мозга найти невозможно. Следовательно, приходится предполагать, что в активности мозга существенную роль играют, как минимум, обычные квантовые процессы (если не квантово-гравитационные, как предполагает Пенроуз).

Коль скоро мозг существенным образом использует квантовые процессы для формирования и обработки смыслов, то и смыслы, о которых писал Дрейфус, по крайней мере частично, связаны с квантовыми состояниями, и не имеют простой информационной природы. Скорее всего, смыслы могут иметь смешанную информационно-квантовоинформационную природу. Компьютеры обрабатывают только информацию в обычном понимании, поэтому мозг, обрабатывая и квантовую информацию, уже только по этой причине не является обычным компьютером. Квантовые состояния нельзя копировать, но можно переносить в единственном экземпляре с одной системы на

другую похожую систему (с той же размерностью гильбертова пространства состояний). Такой перенос можно осуществить с помощью процесса квантовой телепортации [27, с. 309-315]. В частности, если в основе феномена сознания лежат квантовые процессы, то состояние сознания принципиально невозможно «скопировать» на искусственный носитель, оставив «оригинал» сознания в целости и сохранности.

Парадокс квантовой информации и «чудо клонирования». Действительно, в реальном мире квантовое состояние нельзя скопировать никакими силами. Однако, предположим, мы создали точную вычислительную модель некоторой квантовой системы, включая симуляцию физического пространства, в котором «живет» квантовая система. Это возможно в силу вычислимости квантовой теории. Как уже говорилось, квантовая система и ее вычислительная модель в классическом компьютере, это, с точностью до изоморфизма, одно и то же. Но в вычислительной модели, в отличие от реальной квантовой системы, квантовые состояния уже имеют ясную информационную природу. Это просто последова-тельности чисел, кодированные битами компьютерной памяти, которые представляют векторы гильбертова пространства. Очевидным образом, можно создать сколько угодно копий такого объекта (хотя копии невозможно будет создать только средствами, формализованными внутри модели). Почему в компьютерной симуляции квантовое сос-тояние представляет собой информацию, а в оригинале квантовой системы - нет? Не противоречит ли обычная информационная природа симулированного квантового состояния неинформационной природе реальных квантовых состояний?

Нам представляется, что противоречие, действительно, есть, и оно очень серьезно. По нашему мнению, наличие этого противоречия означает, что оригинальная

квантовая система и ее даже очень совершенная вычислительная модель на классическом компьютере, - в действительности не одно и то же с точностью до изоморфизма.

Мы ясно понимаем, что возможен также иной взгляд на вещи. Именно, можно заключить, что информационная природа компьютерной модели квантового состояния указывает на то, что и реальные квантовые состояния имеют информационную природу, несмотря на существование теоремы о запрете клонирования.

Можно также заметить, что квантовое состояние и симуляция квантового состояния классическим компьютером представляют разные типы реальности. Как подробно обсуждается в нашей статье [24], достаточным критерием объективного существования некоторого объекта является возможность получения о нем воспроизводимых данных воспроизводимыми методами. Одиночное квантовое состояние не удовлетворяет этому критерию: воспроизводимая методика измерений (например, измерение спина в установке Штерна-Герлаха) приводит, вообще говоря, к невоспроизводимым результатам (непредсказуемое направление проекции спина). Напротив, симуляция квантового состояния на классическом компьютере вполне удовлетворяет критерию объективной реальности. В отличие от одиночного квантового состояния, критерию объективной реальности удовлетворяет также ансамбль систем, представляющий квантовое состояние. Для ансамбля состояний воспроизводимые измерения приводят к воспроизводимым результатам в отношении различных распределений вероятности, а использование методов квантовой томографии позволяет полностью восстановить квантовое состояние, соответствующее ансамблю. Восстановив же это состояние, его можно скопировать на состояние другой квантовой системы с той же размерностью гильбертова пространства состояний или просто записать на классический носитель информации для последующего использования. То есть, в отличие от одиночного квантового состояния, ан-

самбль квантовых состояний является информационным понятием в том смысле, что в нем может быть закодирована информация в обычном смысле.

Парадоксальное противоречие между неинформационной природой квантового состояния и информационной природой симуляции квантового состояния можно превратить в еще более сильный парадокс. Предположим, что на классическом компьютере удалось очень точно смоделировать настолько большой фрагмент квантовой реальности, что он может включать в себя разумного наблюдателя. Если симуляция достаточно точна, то такой виртуальный наблюдатель не сможет заметить, что он живет в виртуальном, а не в реальном мире. Но теперь мы возьмем, и в нашей компьютерной симуляции начнем создавать копии состояний некоторых квантовых систем, которые наблюдатель способен исследовать. Для нас («программистов») здесь нет никакой проблемы, так как в компьютерной модели квантовые состояния — это просто классические цепочки бит. Однако наблюдателю известна теорема о запрете клонирования состояний, и он должен будет расценить «феномен клонирования», происходящий у него а глазах, как чудо, нарушение законов природы.

Космологический горизонт вычислимости. Возможность разрешения парадокса «чуда клонирования» кроется в критическом анализе понятия вычислимости по отношению к квантовой теории. Фактически, для вычисления поведения некоторых практически важных квантовых систем требуются такие мощности классического компьютера или такие объемы вычислений, которые нельзя реализовать не только практически, но и принципиально, так как для вычислений будет необходимо время, превышающее возраст Вселенной, или размер компьютера будет таков, что его нельзя будет уместить внутри космологического горизонта событий. Вот простейший пример. Для того, чтобы с использованием квантового алгоритма Шора разложить на простые

множители 1000-значное двоичное число (обычная задача для перспективных квантовых компьютеров), квантовому компьютеру требуется память всего в тысячу квантовых ячеек памяти - кубитов, в то время как классическому компьютеру для представления состояния такого квантового компьютера потребуется память $2^{1000} \approx 10^{300}$ комплексных чисел. Это на много порядков больше объема информации, которая может быть записана во всем обычном веществе видимой части Вселенной ($\sim 10^{90}$ бит, [32]). Даже если внутри видимого космологического горизонта попытаться разместить по одному биту информации в каждой планковской ячейке пространства (объемом порядка 10^{-99} см^3), то всего поместится порядка 10^{183} бит, что все еще почти на 120 порядков меньше необхо-димого. Квантовую систему, отвечающую 1000-кубитному квантовому компьютеру принци-пиально невозможно симулировать на классическом компьютере, поэтому, в частности, реальные классические компьютерные симуляторы квантовых вычислений могут работать только с очень малоразмерными системами.

Этот пример представляет очень общую проблему. Точное решение практически любой нетривиальной многочастичной проблемы в квантовой теории требует сверхкосмо-логически больших вычислительных ресурсов. По этой причине, например, принциально неразрешима проблема точного вычисления спектров возбуждений атомных ядер, состоящих больше чем из нескольких нуклонов и т.д.

Подчеркнем, что ограничение на возможность симуляции сложных (многочастичных) квантовых систем имеет принципиальный характер. Это ограничение связано, как мы видели, с фундаментальными характеристиками Вселенной, в которой мы обитаем (космологический горизонт событий, связанный с конечным возрастом Вселенной), и его уместно назвать *космологическим горизонтом вычислимости.* Он столь же фундаментален и непреодолим, как космологический горизонт событий. Понятие вы-

числимости в отношении сложных многочастичных квантовых систем лишено физического смысла, если необходимые ресурсы классического компьютера выходят за космологический горизонт вычислимости. Такие системы не могут быть исчерпывающим образом симулированы классическим компьютером в нашей Вселенной. Но, поскольку нам доступна единственная Вселенная, следует считать, что они вовсе не допускают симуляции классическими средствами, не являются вычислимыми в смысле космологического горизонта. Такая невычислимость имеет статус закона природы.

Понятие космологического горизонта вычислимости позволяет снять парадокс «чуда клонирования». Любые достаточно большие фрагменты квантовой реальности, очевидным образом, невычислимы в смысле космологического горизонта.

Космологический горизонт вычислимости, перспективы симуляции мышления и природа теоремы Пенроуза. Если работа мозга на некотором фундаментальном уровне опирается на что-то похожее на квантовые вычисления, то размерность соответствующей квантовой системы (или систем) будет столь высока, что ее физика заведомо будет невычислимой в смысле космологического горизонта вычислимости. Говорить о вычислимости поведения квантового мозга физически бессмысленно. Однако, физика мозга все еще сохраняет свою алгоритмически разрешимую (вычислимую) природу в чисто математическом смысле. Вопрос состоит в том, должны ли мы считать квантовую активность мозга вычислимой, или нет, если она невычислима с физической, но вычислима с математической точки зрения?

С нашей точки зрения, этот вопрос настолько сложен, что следует признать, что на него пока невозможно дать вполне определенный ответ. В то же время, идея Роджера Пенроуза искать природу невычислимой активности мозга в еще не известной физике, основана на определенном от-

вете *«да»* на этот вопрос. Представляется, что такая определенность является несколько преждевременной, поэтому возможность симулировать мышление человека не выходя за пределы квантовых вычислений не кажется закрытой.

Понятие космологического горизонта вычислимости позволяет несколько по-новому взглянуть на саму теорему Пенроуза. Доказательство Пенроуза апеллирует к теореме Гёделя-Тьюринга, а теорема Гёделя-Тьюринга использует представление о завершающихся и незавершающихся алгоритмах. Когда речь идет о завершающихся алгоритмах, то нигде в теореме, конечно, не затрагивается вопрос о том, как быстро они завершаются. Важна только «принципиальная» завершаемость. Роджер Пенроуз поясняет эту мысль следующим примером. На стр. 141 в книге [3] он рассматривает следующую задачу для компьютера «распечатать последовательность единиц, после чего остановиться». Пенроуз считает, что такую задачу следует считать завершающейся (вычислимой), несмотря на то, что необходимое время на такое вычисление превышает космологический горизонт вычислимости. Поэтому теорема Гёделя-Тьюринга, а следовательно и теорема Пенроуза, явно оперируют вычислимостью и невычислимостью чисто математического типа, независимого от космологического горизонта вычислимости.

Однако, реальность может быть такова, что Гёделевское утверждение мы принципиально не можем «иметь в руках» по той причине, что возможности генерации этого утверждения окажутся за пределами космологического горизонта вычислимости. Может быть, это и не произойдет, но доказательство теоремы нам ничего в этом смысле не гарантирует. Не означает ли это, что уже теорема Гёделя-Тьюринга лишена физического смысла? И допустимо ли понятие физического смысла распространять на математические теоремы? Но ведь вопрос о создании сильного ИИ - это практический, а не абстрактно-математический вопрос. Мы не готовы ответить на эти вопросы, они требуют даль-

нейших размышлений. Как нам представляется, самое сложное в этих вопросах - это понять, какой методикой надо вооружиться для ответа на них.

Резюме

Очевидным выводом из аргументации, приведенной в данной статье, является то, что основным препятствием на пути создания сильного ИИ является вовсе не недостаточная мощность современных компьютеров, а множество принципиальных нерешенных проблем, не все из которых даже удается аккуратно сформулировать. По мнению автора, число этих проблем столь велико, и они настолько трудны, что создание сильного ИИ в ближайшие десятилетия представляется очень маловероятным.

Что касается теоремы Пенроуза об искусственном интеллекте, то при всей ее невероятной силе, которую автор признает и всячески старался подчеркивать в статье, в ней остаются тонкие и не вполне ясные места. Интересно и полезно то, что анализ теоремы стимулирует новый взгляд на старые проблемы.

Список литературы

1. Kurzweil R. The singularity is near: when humans transcend biology. – USA: Viking Penguin, 2005.
2. Пенроуз Р. Новый ум короля. – М.: УРСС, 2003.
3. Пенроуз Р. Тени разума: В поисках науки о сознании. – Москва-Ижевск: Институт компьютерных исследований, 2005.
4. Strong AI (статья в Википедии) / http://en.wikipedia.org/wiki/Strong_AI
5. Searle J. Minds, Brains and Programs // Behavioral and Brain Sciences. 1980. Т. 3. № 3. – С. 417–427.
6. Vernor Vinge V. The Coming Technological Singularity: How to Survive in the Post-Hu-man Era // VISION – 21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute, March 30-31 – 1993.

http://www-rohan.sdsu.edu/faculty/vinge/misc/singularity.html
7. Joy B. Why the future doesn't need us // Wired – April 2000 – P. 238–262.
8. Forster H., Mora P., Amiot L.W. Doomsday: Friday, 13, November, A.D. 2026 // Science. 1960 V.132 – P. 1291–1295.
9. Шкловский И.С. Вселенная, жизнь, разум. – М.: Наука, 1965.
10. Робот-учёный делает открытия без помощи человека / http://www.infuture.ru/article/1917
11. Редько В.Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики. – М.: Книжный дом «ЛИБРОКОМ», 2013.
12. Дунин-Барковский В.Л. К вопросу об обратном конструировании мозга // В кн.: «Глобальное будущее 2045. Конвергентные технологии и трансгуманистическая эволюция». Под ред. проф. Д.И. Дубровского. – М.: ООО «Издательство МБА», 2013. - С. 150–157.
13. Boyle J.H. *C. elegans* locomotion: an integrated approach. PhD thesis. The University of Leeds School of Computing, 2009.
14. Слизевики – «разумные» простейшие / http://lostlab.ru/forum/topic432.html
15. Александров В.Я. Проблемы поведения на клеточном уровне - цитоэтология // Успехи современной биологии. – 1970. – Т. 69. –№ 2. – С. 220–240.
16. Панов А.Д. Универсальная эволюция и проблема поиска внеземного разума (SETI). - М.: ЛКИ (URSS), 2008.
17. Режабек Б.Г. О поведении механорецепторного нейрона в условиях замыкания его цепью искусственной обратной связи // ДАН СССР. – 1971. – Т. 198. – № 4. – С. 981–984.
18. Penrose R., Hameroff S. Consciousness in the Universe: Neuroscience, Quantum Space Time Geometry and Orch OR Theory // Journal of Cosmology, 2011. Vol. 14. http://journalofcosmology.com/Consciousness160.html
19. Цитоскелет (статья в Википедии) / http://ru.wikipedia.org/wiki/Цитоскелет

20. Craddock T.J.A., Tuszynski J.A., Hameroff S. Tuszynski Cytoskeletal Signaling: Is Me-mory Encoded in Microtubule Lattices by CaMKII Phosphorylation? // PLOS Comput Biol. 2012. V. 8(3).
21. van der Sar T., Wang Z.H., Blok M.S., Bernien H., Taminiau T.H., Toyli D.M., Lidar D.A., Awschalom D.D., Hanson R., Dobrovitski V.V. // Nature. 2012. V. 484. – P. 82–86.
22. Фейнман Р., Лейнтон Р., Сэндс М. Фейнмановские лекции по физике. Т.8. - М.: Мир, 1966.
23. Клини С.К. Введение в метаматематику. - М.: Изд-во ИЛ, 1957.
24. Панов А.Д. Природа математики, космология и структура реальности: объективность мира математических форм // В кн.: Космология, физика, культура, М.: ИФРАН, 2011. – С. 191-219.
25. Дрейфус Х. Чего не могут вычислительные машины. – М.: ЛИБРОКОМ (URSS), 2010.
26. Менский М.Б.. Концепция сознания в контексте квантовой механики // Успехи физических наук. 2005. Т. 175. № 4. – С. 413–435.
27. Менский М.Б. Человек и квантовый мир. – Фрязино, Век 2, 2007.
28. Менский М.Б. Сознание и квантовая механика. Жизнь в параллельных мирах (Чудеса сознания - из квантовой реальности). – Век 2, Фрязино, 2011.
29. Линде А.Д. Интервью для журнала «Вокруг Света» (Сергей Добрынин) // Вокруг Света. 2012. № 10 (2865). – С. 139–148.
30. Julia-Dıaz B., Burdis J.M., Tabakin F. Qdensity – a Mathematica quantum computer simulation // arXiv:quant-ph/0508101, 2005.
31. Гриб А.А. Неравенства Белла и экспериментальная проверка квантовых корреляций на макроскопических расстояниях // Успехи физических наук. 1984. Т. 142. – С. 619-634.
32. Гуревич И.М., Урсул У.А. Информация - всеобщее

свойство материи: Характеристики, оценки, ограничения. – М.: Книжный дом «ЛИБРОКОМ» (URSS), 2012.

УДК 519.25
А.К. Горшенин, канд. физ.-мат. наук, ст. научн. сотр.
С.Я. Шоргин, д-р физ.-мат. наук, зам. дир.
Институт проблем информатики РАН, г. Москва, Россия
agorshenin@ipiran.ru

РАЗРАБОТКА ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА БОЛЬШИХ ДАННЫХ

Введение

В статье существенное внимание уделяется работе с большими данными произвольной природы, рассматриваются возникающие задачи и актуальность их решения. На примере анализа реальных данных из области финансовых рынков и физики турбулентной плазмы формулируются возможные подходы, которые могут позволить существенно уточнить тонкую структуру сложных процессов с помощью статистических моделей. Данные модели составляют основу информационной технологии, позволяющей автоматически обрабатывать поступающие входные данные из различных областей.

Большие данные

Термин «большие данные» обычно обозначает объединение огромных по размеру объемов данных и, одновременно, соответствующих методов хранения и обработки. При этом обычно считается, что объем данных настолько большой, а их структура столь сложна, что применение традиционных техник работы представляется малоэффективным, а чаще и вовсе невозможным.

Ключевая проблема в отношении ресурсов и данных представлена на рисунке 1 (источник IDC): количество производимой информации в мире, контента, становится больше с каждой минутой, в то время как ресурсы (физи-

ческие, вычислительные) на практике обычно ограничены различными факторами. На графике видно разделение двух кривых, начиная с 2007 года, при этом рост информационных объемов имеет экспоненциальный характер – и уже в 2011 году разрыв был существенен. Отметим, что по оси ординат отложены значения в экзобайтах.

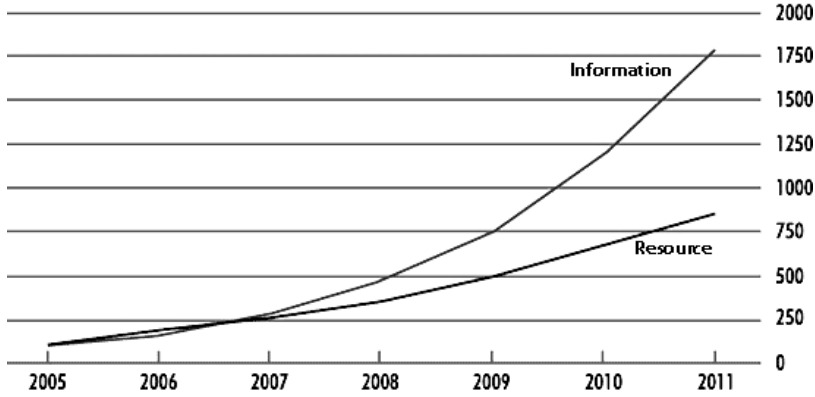


Рисунок 1 – Объем создаваемой информации и доступные ресурсы хранения данных

Обычно для больших данных формулируется так называемый принцип четырех «V», в котором первые буквы ключевых характеристик больших данных начинаются с соответствующей латинской буквы. Приведем классификацию, предлагаемую компанией IBM [1].

1. Объем (volume): компаниям, научным центрам, предприятиям приходится каждый день сталкиваться с нарастающими объемами информации, которые могут измеряться тера- и даже петабайтами.

2. Скорость (velocity): существует огромное количество областей, обработка данных в которых должна производиться в реальном времени вне зависимости от объемов поступающей информации. В бизнес-сфере стоит упомянуть финансовые системы, в которых мошеннические действия должны выявляться крайне быстро. В противном случае, финансовые потери организации (и ее клиентов) сложно прогнозировать.

3. Разнообразие (variety):большие данные не являются данным только одного определенного типа. Это могут быть и текстовые файлы, и мультимедийные файлы, и графики, и показатели приборов. Области формирования больших данных огромны – речь о них пойдет ниже.

4. Достоверность (veracity): в бизнесе, при проведении научных экспериментов крайне важна предсказуемость собственных действий. Поэтому крайне важно доверять результатам проведенных исследований – а для этого они должны быть достоверны.

Теперь рассмотрим области, являющиеся ключевыми для обработки больших данных. Прежде всего, начнем с «Big science», то есть научных исследований, которые являются крайне дорогостоящими и требуют привлечения значительного числа научно-исследова-тельских групп. Безусловным примером в данной области является Большой адронный коллайдер (БАК). Сейчас в день объем поставляемых для обработки данных с БАК составляет 500 экзабайт, однако в самое ближайшее время ожидается увеличение до 500 квинтиллионов байт, то есть увеличение объема данных на два порядка. Стоит заметить, что такой объем почти в 200 раз превышает размер информации, поставляемой со всех других источников в мире! Безусловно, БАК является крайне большой и важной с точки зрения науки системой, однако стоит отметить, что к вводу в эксплуатацию готовится токамак, который будет поставлять информацию также экзобайтами. Очевидно, число таких систем в ближайшее время будет только нарастать.

Помимо огромных научных систем всемирного масштаба, существует значительное число систем меньшего размера, также поставляющих значительные объемы научной информации, например, космические телескопы, системы прогноза погоды, системы моделирования экспериментов.

О важности работы с большими данными говорит и тот факт, что в разных европейских странах и США созда-

ются официальные комиссии по стандартизации понятия «big data», а также методов обработки. В качестве примера достаточно упомянуть так называемую «Big Data Research and Development Initiative» [2]. При этом ведущие ИТ-компании (IBM, Microsoft, HP, Google и др.) уже предлагают некий набор решений для работы с большими данными – и в тоже время ищут специалистов в этой области (data scientist), хотя выпускников непосредственно с такими дипломами найти достаточно сложно.

Наконец, стоит упомянуть о частных компаниях. Примеры здесь могут быть весьма разнообразными. В качестве небольшой справки приведем несколько показателей для популярных социальных сетей. Так, в Twitter ежедневно пользователи создают 12 терабайт контента. В сети Facebook пользователи хранят более 50 миллиардов фотографий – и их число увеличивается с каждым днем.

Совершенно очевидно, что в данных условиях создание и развитие методов хранения, обработки гигантских объемов информации становится не просто важной, но критической задачей, решение которой направлено как на повышение эффективности бизнеса, так и на получение принципиально новых научных результатов в различных областях.

Метод скользящего разделения смесей

Рассмотрим общую концепцию анализа больших данных достаточно общего типа, а именно информационных потоков. Действительно, существует большой набор областей, в которых поступление данных вполне может соотноситься с некоторыми событиями – что автоматически и формирует информационный поток. При этом мы используем базовое предположение о стохастическом характере данных, предполагая, что они могут быть промоделированы смесью некоторых вероятностных распределений.

Наилучшей моделью однородного хаотического потока событий является пуассоновский процесс. Однако в

реальных «хаотических» системах хаос практически никогда не бывает однородным в пространстве или времени. По-видимому, это связано с тем, что ни одна из реально функционирующих сложных информационных систем не может рассматриваться как замкнутая, то есть изолированная от окружающей среды. Наиболее разумными стохастическими моделями неоднородных хаотических точечных процессов являются дважды стохастические пуассоновские процессы, иначе называемые процессами Кокса (см., например, книгу [3]). В связи с указанными обстоятельствами в качестве разумной модели представляется использование конечной смеси гамма-распределений, то есть распределения с плотностью вида

$$\sum_{i=1}^{n} p_i \cdot f_{r_i,\theta_i}(x),$$

где через

$$f_{r_i,\theta_i}(x) = \frac{\theta_i^{r_i} x^{r_i-1} e^{-\theta_i x}}{\Gamma(r_i)}$$

обозначена стандартная плотность гамма-распределения, величины p_i называются весами и в сумме равны единице.

Теперь воспользуемся так называемым методом скользящего разделения смесей (СРС-методом [3]). Основной принцип этого подхода заключается в том, что мы разбиваем исходную выборку на некоторые подвыборки, обычно существенно меньшего размера, называемые окнами, и проводим анализ состояния системы на каждом из этих окон. Далее, окно сдвигается на несколько позиций (элементов) – и анализ производится заново. Отметим, что величина этого сдвига может меняться – и на соседних шагах мы можем анализировать как выборки, отличающиеся всего на один элемент, так и полностью непересекающиеся части исходного ряда. Данный подход позволяет отследить динамическое изменение состояния системы во времени и получить вид эволюции формирующих процессов в ней.

В качестве приближения параметров смеси будем использовать оценки максимального правдоподобия, которые можно найти с помощью EM-алгоритма [3], реализованного для гамма-распределений. Формулы для вычисления параметров гамма-распределения на итерационных шагах имеют следующий вид. Параметр формы может быть найден из численного решения уравнения

$$\log r_i^{(m)} - \psi(r_i^{(m)}) = -\frac{\sum_{j=1}^{n} g_{ij}^{(m)} \log \frac{x_j}{A_i^{(m)}}}{\sum_{j=1}^{n} g_{ij}^{(m)}},$$

где $\psi(\cdot)$ – дигамма-функция, величина $A_i^{(m)}$ имеет вид

$$A_i^{(m)} = \frac{\sum_{j=1}^{n} x_j g_{ij}^{(m)}}{\sum_{j=1}^{n} g_{ij}^{(m)}}, \quad g_{ij}^{(m)} = \frac{p_i^{(m)} f_{r_i^{(m)},\theta_i^{(m)}}(x_j)}{\sum_{l=1}^{k} p_l^{(m)} f_{r_l^{(m)},\theta_l^{(m)}}(x_j)},$$

параметр масштаба определяется на каждом шаге из соотношения

$$\theta_i^{(m)} = \frac{A_i^{(m)}}{r_i^{(m)}}, \; i = 1, \ldots, k,$$

веса задаются формулой

$$p_i^{(m+1)} = \frac{1}{n} \sum_{j=1}^{n} g_{ij}^{(m)}.$$

Данный подход позволяет выявить как основные процессы, оказывающие существенное влияние на функционирование системы, так шумовые процессы, которые могут возникать по различным причинам, в том числе, демонстрируя вычислительные погрешности.

Примеры применения к реальным данным

В настоящем разделе рассмотрим применение методологии для данных из весьма различающихся между собой предметных областей. В качестве первого примера рассмотрим так называемую биржевую книгу заявок.

Для более точного объяснения данного понятия кратко рассмотрим механизм торгов на фондовой бирже. Трейдеры (инвесторы) выставляют заявки (делают поручения брокеру) на продажу либо приобретение различных инвестиционных инструментов (ценных бумаг и т.п.) по выбранной стоимости. После выставления заявки на бирже производится автоматическая проверка всех встречных заявок, имеющих указанную или лучшую цену (таким образом, к примеру, для продажи ищутся встречные заявки с ценой равной или большей указанной в заявке). Сделка происходит, когда такая встречная заявка будет найдена. Если подходящих заявок нет, то выставленная заявка останется на бирже и будет ожидать появления подходящих встречных заявок (также она может быть отменена заявителем). В течение торгов в каждый момент времени в базе данных биржи находится очередь неисполненных заявок на продажу и на покупку. Этими заявками и образуется биржевая книга заявок – совокупность заявок на покупку и на продажу, информация о которых отражается в виде очереди в торговом терминале. Данные такого рода описывают функционирование реальной сложной информационной финансовой системы электронных торгов. Если в качестве события рассматривать любое изменение в биржевой книги, то с учетом дробления по временным интервалам, получается более ста тысяч элементов для анализа ежедневно!

На рисунке 2 представлены графические результаты, а именно, эволюция плотности распределения событий в биржевой книге заявок на верхнем графике, и одно из сечений плотности с соответствующей гистограммой на нижнем графике. На графике для эволюции плотности на

осях присутствуют обозначения «Время» (для каждого сечения), «Номер окна» (обозначает положение нашего окна в данных), а также «Значение» для графика плотности. Цветовая шкала в данном случае отображает разницу между маленькими и большими величинами (от глубокого синего к темно-красному). Нижний график показывает «Значение» и «Время», демонстрируя, во-первых, тот факт, что обычное гамма-распределение не вполне адекватно будет приближать исходные данные, а во-вторых, можно увидеть качество приближения гистограммы нашей смоделированной кривой.

В качестве второго примера рассмотрим эволюцию во времени спектров, полученных в одном из режимов функционирования турбулентной плазмы. А именно, обрабатывались данные с диагностики рассеяния излучения на второй гармонике гиротронов. Данная диагностика измеряет коротковолновые флуктуации плазмы вблизи центра плазменного шнура. Исходные данные представляют собой односторонний несимметричный Фурье-спектр, убывающий с увеличением частоты. Разложение спектра на компоненты в данном случае дает представление о поведении различных типов колебаний, существующих и эволюционирующих во времени в плазме. Для анализа используется метод бутстреп-моделирования выборки, поэтому для каждого спектра, полученного в некий момент времени, можно формировать выборки значительного объема (например, сто тысяч элементов) – и поступать так для каждого временного интервала.

На рисунке 3 представлены графические результаты эволюци спектра для плазмы. Здесь для эволюции спектров на осях присутствуют обозначения «Частота» (для каждого сечения), «Номер окна» (обозначает положение нашего окна в данных), а также «Значение» для графика плотности. Цветовая шкала в данном случае также показывает разницу между маленькими и большими величинами (от глубокого синего к темно-красному).

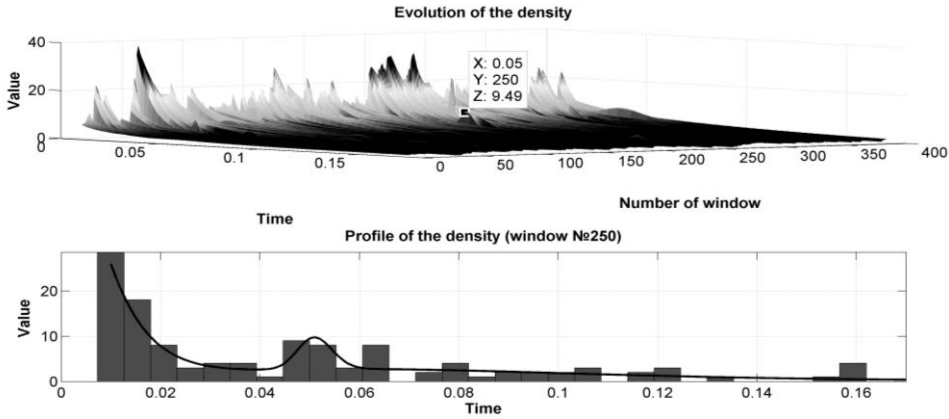


Рисунок 2 – Эволюция плотности для биржевой книги заявок

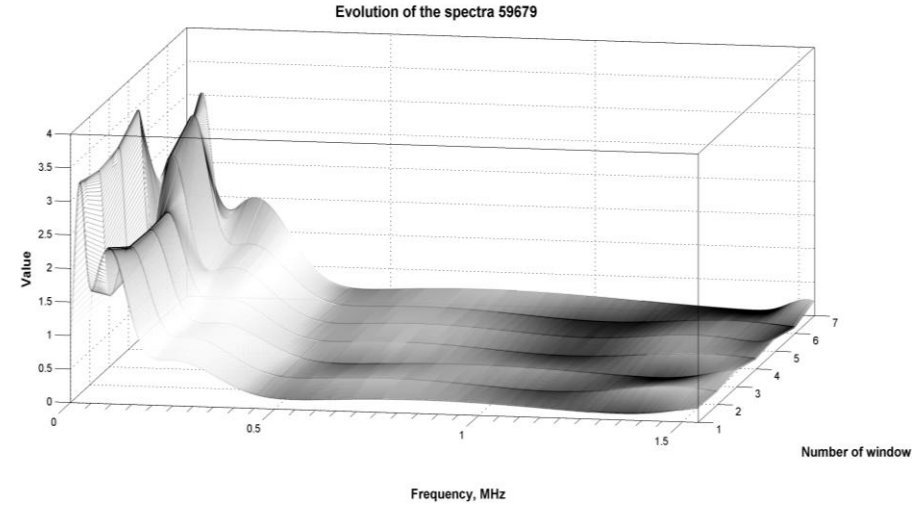


Рисунок 3 – Эволюция спектра для плазмы

В частности, проведенные исследования позволили установить ряд важных с физической точки зрения результатов: например, было обнаружено, что выделяются три доминирующие компоненты, у которых меняется пропорциональный состав, что соответствует перекачке энергии между турбулентностями различного типа. Алгоритмы по-

зволили выявить недостатки при получении экспериментальных данных, связанные с погрешностями приборов и т.п.

Выводы

Большинство специалистов сходится во мнении, что технология big data может предоставить конкурентные преимущества тем, кто найдет новые, оригинальные способы обработки данных. Таким образом, данное направление становится ключевым практически во всех сферах в относительно долгосрочной перспективе, так как однозначно эффективных методов, а также их строгих теоретических обоснований на сегодняшний день не существует.

В представленной работе показаны результаты, позволяющие судить о качестве анализа данных с помощью предложенной методологии. Дополнительные результаты, связанные с анализом информационных потоков «в чистом виде» (связанные с анализом траффика) на основании данного подхода, представлены в работе [4].

Безусловно, при работе со столь большими объемами данных крайне важным фактором становится скорость обработки. Однако, оставаясь в рамках предложенной методологии, возможны некоторые ускорения алгоритмов без принципиального изменения их сути (см., например, работу [5], посвященную сеточным методам, которые могут быть весьма эффективными в ряде ситуаций). Таким образом, открываются возможности для построения вполне удобных инструментов работы с большими данными.

Авторы выражают глубокую признательность профессору В.Ю. Королеву. за полезные обсуждения при проведении исследований.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проекты 12-07-00115а, 12-07-31267мол_а и 11-07-00112а).

Список литературы

1. What is big data? [Электронный ресурс]. – http://www-01.ibm.com/software/data/bigdata
2. Big Data Initiative–The White House [Электронный ресурс]. – http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf 29 March 2012.
3. Королев В.Ю. Вероятностно-статистические методы декомпозиции волатильности хаотических процессов. – М.: изд-во Моск. ун-та, 2011. – 512 с.
4. Gorshenin A., Korolev V. Modeling Of Statistical Fluctuations Of Information Flows By Mixtures Of Gamma Distributions / A. Gorshenin, V. Korolev // Proceedings 27th European Conference on Modelling and Simulation – 2013, Alesund, May 27-30, 2013. - С. 569 –572.
5. Gorshenin A., Korolev V., Kuzmin V., Zeifman A. Coordinate-Wise Versions Of The Grid Method For The Analysis Of Intensities Of Non-Stationary Information Flows By Moving Separation Of Mixtures Of Gamma-Distribution / A. Gorshenin, V. Korolev, V. Kuzmin, A. Zeifman // Proceedings 27th European Conference on Modelling and Simulation – 2013, Alesund, May 27-30, 2013. - С. 565 –568.

УДК 621
И.Н. Синицын, д-р техн. наук, проф.
А.С. Шаламов, д-р техн. наук, проф.
В.И. Синицын, д-р физ-мат. наук
Э.Р. Корепанов, канд. техн. наук
В.В Белоусов, канд. техн. наук
А.А.Кулешов, аспирант
Институт проблем информатики РАН, г. Москва, Россия
nordxsi@gmail.com

МЕТОДЫ И СРЕДСТВА ОЦЕНКИ ЗАПАСОВ И УРОВНЯ ГОТОВНОСТИ СИСТЕМ ИНТЕГРИРОВАННОЙ ЛОГИСТИЧЕСКОЙ ПОДДЕРЖКИ, ОСНОВАННЫЕ НА КАНОНИЧЕСКИХ РАЗЛОЖЕНИЯХ СЛУЧАЙНЫХ ФУНКЦИЙ

Введение

Как известно [1], современные подходы к использованию информационно - вычислительных систем в сфере управления стоимостью жизненного цикла изделий наукоемкой продукции (ЖЦ ИНП) основаны на CALS – методологии (Continuous Acquisition and Life cycle Support - непрерывная информационная поддержка поставок и жизненного цикла), направленной на достижение эффективных результатов в указанной области. Основу этой концепции составляют технологии интегрированной логистической поддержки (ИЛП), средства и процедуры которой должны обеспечивать минимизацию затрат покупателя на послепродажное сопровождние. С этой целью создаются специализированные технико-экономи-ческие информационные модели ЖЦ ИНП, представляющие собой нормативные базы данных и технической документации. Эти модели носят статический характер, что не позволяет оперативно управлять качеством и стоимостью на заданном периоде эксплуатации изделий. Поэтому дополнительно использу-

ют базы данных мониторинга основных эксплуатационных и других необходимых характеристик изделия, позволяющих вычислять главные показатели эффективности управления, включая стоимость и технологичность эксплуатации.

Однако при всем совершенстве и многообразии технологий ИЛП, предлагаемых этими стандартами, фактически остается нерешенной главная проблема – проблема оптимального проектирования системы обслуживания, ремонта и снабжения для обеспечения эксплуатации изделий с минимальной стоимостью на годы вперед, вплоть до их списания.

Дается обзор работ, выполненных в ИПИ РАН в области теории и практики систем интегрированной логистической поддержки (СИЛП) изделий наукоемкой продукции.

Приводятся результаты разработки и испытаний типовой инструментальной системы «Динамика процессов» («ASOS») версия 2.0, предназначенной для автоматического поиска и анализа оптимальных поставок запасных частей в системах послепродажного обслуживания изделий наукоемкой продукции. Они позволяют кардинально реформировать традиционные системы управления при создании и эксплуатации ИНП путем внедрения методов оптимального планирования процессов расходования ресурсов, по критериям экономической целесообразности и эффективности.

Развитие методов анализа и синтеза СИЛП, основанное на канонических разложениях с независимыми компонентами

В [1, 2, 3-7] подробно представлены математические модели количественного описания различных ресурсов в СИЛП. Они представляют собой нелинейные гибридные стохастические системы (СтС) с параметрическими шумами. Корреляционные алгоритмы моделирования и прогно-

зирования для гибридных СтС основаны на детерминированных уравнениях для математических ожиданий, а также матрицы ковариационных функций переменных состояния Y_i и начальных условий поставок $\pi_i(t_j)$. Были разработаны специальные методы эквивалентной статистической линеаризации для гибридных СтС Ито.

На основе методов теории канонических разложений с независимыми компонентами (КРНК) случайных функций [8, 9] для СтС большой размерности разработаны эффективные методы аналитического моделирования и параметрического синтеза СИЛП. При использовании КРНК:

$$Y(t) = m^y(t) + \sum_{h=1}^{H} V_h y_h(t), \qquad (1)$$

($m^y(t)$ - вектор математического ожидания, V_h - независимые скалярные случайные величины, $y_h(t)$ - векторные детерминированные координатные функции) в рамках корреляционной теории возможны два подхода.

Первый основан, во-первых, на эквивалентной статистической линеаризации исходных соотношений, определяющих нелинейный оператор СИЛП посредством КРНК, во-вторых, составлении нелинейного операторного уравнения для m^y, в третьих, на составлении линейных операторных уравнений для $y_h(t)$, в четвертых, на выписывании начальных условий для m^y и $y_h(t)$. Второй подход основан на решении уравнений для вектора m^y, матрицы ковариаций $K^y(t)$ и матрицы ковариационных функций $K^y(t_1, t_2)$, связанных с (1) формулами:

$$K^y(t) = K^y(t_1 = t, t_2 = t), \qquad (2)$$

$$K^y(t_1, t_2) = \left[K^y_{jl}\right] = \sum_{h=1}^{H} D_h y_{hj}(t_1) y_{hl}(t_2), \qquad (3)$$

где D_h - дисперсии V_h. Для параметрического синтеза используются методы КРНК [9].

Оптимальное планирование материальных средств для поддержания заданного уровня технической готовности парка изделий

На основе разработанных методов [1, 2] для прогнозирования процессов смешанной природы в СИЛП возможны различные постановки задачи. Наиболее общей является глобальная оптимизация комплексной системы «эксплуатация ИНП – обслуживание и ремонт – материально-техническое обеспечение» по критериям стоимости при ограничении, накладываемом на коэффициент (уровень) технической готовности ИНП. При этом подсистема эксплуатации ИНП является источником исходной информации (параметры парка ИНП, интенсивность использования по назначению и др.). В части подсистем обслуживания и ремонта оптимизации могут подлежать параметры производительности предприятий, оказывающих услуги по текущему обслуживанию и ремонту ИНП и составных частей (СЧИ), а также по их капитальному ремонту. В подсистеме материально-технического обеспечения, в первую очередь, необходимо оптимизировать параметры поставок запасных частей, комплектующих и оборудования (сроки, объемы).

В основу инструментальной системы положена упрощенная задача оптимизации - минимизация затрат заказчика на поставки запасных частей на годы вперед, вплоть до списания ИНП. Эта задача имеет вполне самостоятельное значение в условиях, в некотором смысле диктуемых обстоятельствами, когда покупатель согласен с параметрами системы предоставления услуг по обслуживанию и ремонту, обусловленными понятными выгодами для поставщика.

Пусть состояние парка однотипных ИНП и их СЧИ описывается типовым графом, представленным на рисунке

1. Вершины графа отображают возможные состояния составных частей изделия: 1 - исправные СЧИ на складе; 2 - исправные СЧИ, эксплуатируемые в составе изделия; 3 - неисправные СЧИ, находящиеся на восстановительном ремонте у потребителя; 4 - СЧИ в состоянии списания; 5 - исправные СЧИ в составе ИНП, находящихся на профилактических работах; 6 - СЧИ в составе ИНП, находящихся на капитальном ремонте; 7 - неисправные СЧИ, накопленные на складе у потребителя для отправки в ремонт; 8 - СЧИ, находящиеся на восстановительном ремонте у поставщика.

В обороте находится определенное количество СЧИ каждого типа. В силу различных случайных факторов все они распределяются случайным образом между указанными выше состояниями. Введем в рассмотрение фазовый вектор системы $Y(t)$, составляющие которого $Y_i(t), i = 1,\ldots,n$ суть количество СЧИ данного типа, находящихся в i–ом состоянии.

Одним из важнейших показателей ИЛП является средний коэффициент технической готовности (исправности) на заданном промежутке времени $[0,T]$, которое применительно к рассматриваемой системе определяется по формуле:

$$\overline{K}_u(T) = \int_0^T \frac{Y_2(\tau) + Y_5(\tau)}{N} d\tau, \qquad (4)$$

где $Y_2(t)$, $Y_5(t)$ – количество исправных СЧИ; N – находящееся в эксплуатации количество СЧИ данного типа, строго согласованное с количеством эксплуатируемых ИНП (с учетом ИНП, находящихся в капитальном ремонте).

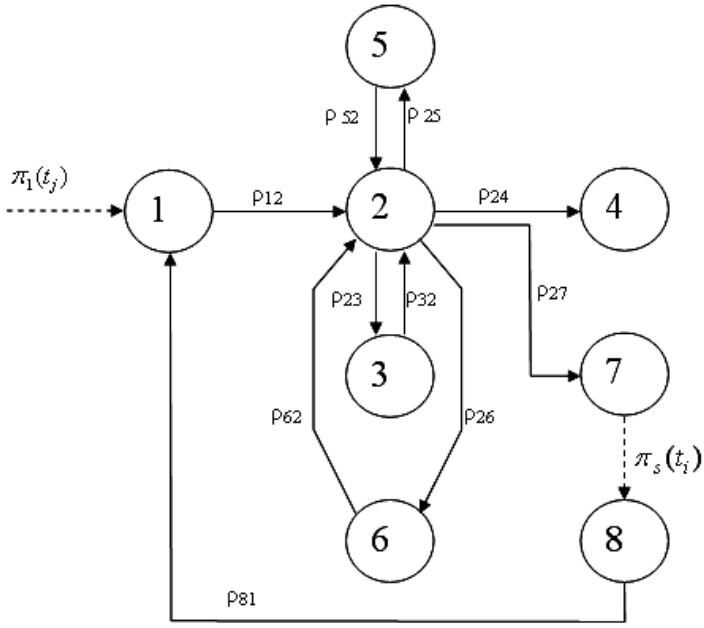


Рисунок 1 - Граф состояний парка изделий и их составных частей

За критерий эффективности (оптимальности) примем

$$J = C_k \int_0^T [\overline{K}_u(\tau) - \overline{K}_u^*(\tau)]^2 \, d\tau, , \qquad (5)$$

где: $\overline{K}_u(\tau)$ – среднее на интервале $[0, t)$ значение коэффициента технической готовности, случайная функция; $\overline{K}_u^*(\tau)$ – заданное заказчиком (требуемое) значение среднего на периоде $[0, T]$ коэффициента технической готовности парка ИНП; C_k – коэффициент нормализации.

Требуется получить оптимальную программу поставок запасных частей (ЗЧ) на заданный период эксплуатации ИНП для обеспечения заданного уровня технической готовности изделий.

Данную задачу решает инструментальный про-

граммный комплекс «Динамика процессов» («ASOS») для прогнозирования и оптимизации, в состав которого входят:

- нормативная база данных;
- база данных мониторинга основных эксплуатационно-технических характеристик ИНП;
- программа прогнозирования и оптимизации «Динамика процессов» 2.0.

«ASOS» решает задачу оптимизации сначала для каждой СЧИ ИНП по отдельности, а затем для изделия в целом.

Исходная информация на момент планирования:

- данные от заказчика (покупателя): штатное количество ИНП (финальных изделий); количество исправных ИНП на момент планирования; средняя наработка СЧИ в год; количество запасных СЧИ на складе; согласованная периодичность поставки запасных частей; заданное значение $\overline{K}_u^*$; горизонт планирования (лет), продолжительность профилактических работ на ИНП на базе заказчика; количество ИНП, находящихся на капитальном ремонте у производителя; средняя наработка СЧИ на отказ; количество СЧИ, находящихся на ремонте у производителей;
- данные от поставщика: назначенный ресурс СЧИ между профилактическими работами, продолжительность заводского ремонта СЧИ, производительность ремонтных предприятий (количество ремонтируемых СЧИ в год), межремонтный ресурс ИНП, производительность системы капитального ремонта ИНП, стоимость услуг, коэффициент эскалации стоимости поставки и др.

Для нахождения оптимальной программы поставок реализована специальная процедура, использующая метод последовательных приближений в ходе поиска значения поставки на очередной год t_{j+1}. Для получения значения критерия в одной точке t_{j+1} (для одной совокупности параметров) необходимо осуществлять «прогон» системы

уравнений на промежутке $\left[0, t_{j+1}\right]$. При этом задействуются ранее вычисленные оптимальные значения поставок $\pi(t_j)$. Данный алгоритм реализован в новой версии специального инструментального программного обеспечения («Динамика процессов» 2.0), которое позволяет автоматически вычислять оптимальную программу поставок запасных частей.

Обсуждение результатов

В ИПИ РАН разработана нового типа информационная система управления, с использованием стохастических подходов для CALS и соответствующего инструментального программного обеспечения прогнозирования и оптимизации работ. В ходе разработки «Динамика процессов» 2.0 были решены основные следующие проблемы:

- создана нормативная база данных паспортизированных СЧИ, содержащая информацию, необходимую для планирования бюджетных показателей СИЛП (назначенные ресурсы, нормы трудозатрат на ППО, ограничения при транспортировке и многие другие);
- создана оперативная база данных, содержащая информацию по мониторингу реальных данных послепродажного обслуживания (наработку изделий, надежность, остаточные ресурсы, дисциплину поставок ЗЧ, дисциплину ремонта и др.);
- разработана математическая модель СИЛП парка ИНП данного типа и заданной численности;
- разработан комплект типовых программных средств в системе Delphi.

Настоящая версия инструментального программного обеспечения, осуществляет автоматический поиск оптимальной программы поставок по годам, позволяет эффективно решать задачу оптимизации не только для упрощенного примера, но и в случае сложных гибридных СППО для крупных предприятий, эксплуатирующих ИНП.

На рисунках 2 и 3 показан пример 1 с результатами

расчета среднего коэффициента исправности на периоде 10 лет в сравнении с заданным, равным 0,8. Показаны верхняя и нижняя граница возможных значений реального среднего коэффициента исправности. На рисунке 2 график соответствует произвольно введенным данным программы поставок одной запасной части по годам. При этом график прогнозируемого среднего коэффициента исправности проходит выше заданного значения 0,8. Стоимость такой программы составит в среднем $1207. На графике также показаны доверительные границы коэффициента, соответствующие вероятности 0,9.

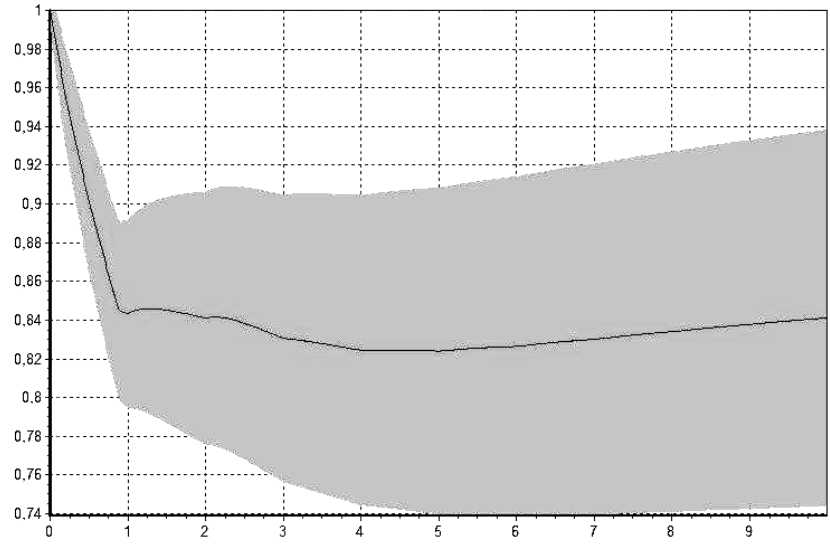


Рисунок 2 - График коэффициента исправности СЧИ, соответствующий произвольной программе поставок

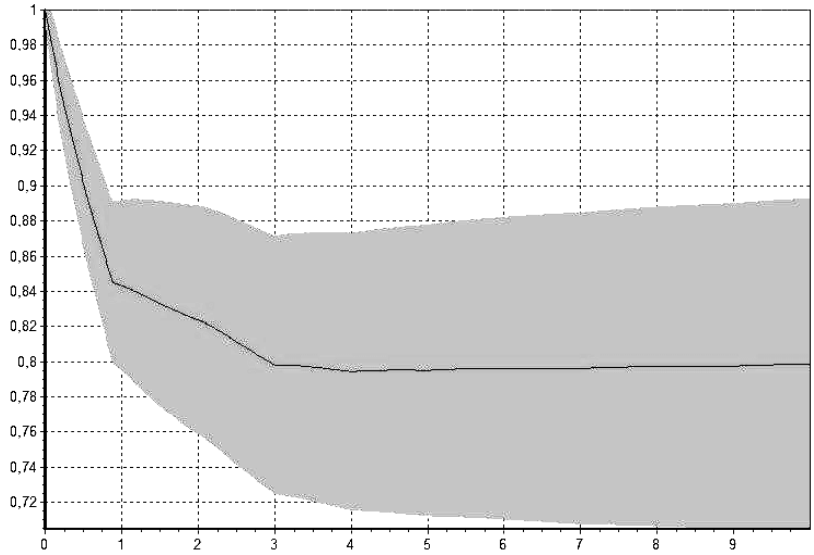


Рисунок 3 - График коэффициента исправности СЧИ, соответствующий оптимальной программе поставок

При использовании последней версии инструментальной системы «Динамика процессов» 2.0, была получена оптимальная программа поставок ЗЧ, при которой средний коэффициент исправности минимально отличается от заданного, равного 0,8 (рисунок 3). Предложенная программа является оптимальной. Стоимость ее составляет $674, что почти в 2 раза меньше ранее предложенной произвольной программы поставок.

Опыт эксплуатации инструмента «Динамика процессов» 2.0 показал следующее:

1. Поставка заказчикам подобных комплексов прогнозирования и оптимизации позволит им формировать оптимальную годовую заявку на поставку запасных частей для проведения как плановых, так и неплановых работ по техническому обслуживанию и ремонту для выполнения плана использования парка ИНП с учетом реальных показателей своей системы ППО, а также на основе данных, сообщаемых поставщиком в соответствии с контрактом.

2. Поставка изготовителям ИНП таких комплексов позволит им формировать управление процессами производства услуг, оценивать перспективы развития производства запасных частей в интересах традиционных заказчиков на годы вперед, а также прогнозировать закупки для перепродажи третьим сторонам.

Версия систем на основе КРНК оказывается достаточно эффективной при размерности СИЛП больше 10. Вместо решения 90 уравнений приходится решать 20 уравнений.

Работа выполнена при поддержке программы ОНИТ РАН «Интеллектуальные информационные технологии, системный анализ и автоматизация» (проект №1.7).

Список литературы

1. Синицын И.Н., Шаламов А.С. Лекции по теории систем интегрированной логистической поддержки. – М.: ТОРУС ПРЕСС, 2012. - 624с.
2. Синицын И.Н., Шаламов А.С. Проектирование CALS систем. Части 1-4 // Системы высокой доступности. М.: Радиотехника. 2012. № 3, 4. 2013. № 1, 2.
3. Синицын И.Н., Шаламов А.С., Сергеев И.В. Проблемы моделирования и минимизации затрат на эксплуатацию изделий наукоемкой продукции на современном этапе // Сб. докладов XIII международной научно-технической конференции «Кибернетика и высокие технологии XXI века». г. Воронеж. Изд. НПФ «Саквоее», 2012. Т.2. - С. 358-370.
4. Синицын И.Н., Шаламов А.С., Корепанов Э.Р., Белоусов В.В., Сергеев И.В., Кулешев А.А. Развитие алгоритмического и инструментального программного обеспечения для аналитического вероятностного моделирования и оптимизации процессов материально-технического обеспечения // Сб. докладов XIV международной научно-технической конференции «Кибернетика и высокие технологии XXI века». г. Воронеж. Изд. НПФ «Саквоее», 2013. - С. 375-384.

5. Синицын И.Н., Шаламов А.С., Сергеев И.В., Синицын В.И., Корепанов Э.Р., Белоусов В.В., Агафонов Е.С., Шоргин В.С. Методы и средств анализа и моделирования стохастических систем интегрированной логистической поддержки // Системы и средства информатики. – М.: ИПИ РАН, 2012. Вып. 22. № 2. - С.3-28.
6. Синицын И.Н., Шаламов А.С., Кулешев А.А. Нелинейное корреляционное моделирование и анализ надежности систем послепродажного обслуживания изделий наукоемкой продукции // Системы и средства информатики. 2013. Т 23. № 1. - С. 80-104.
7. Синицын И.Н., Шаламов А.С., Корепанов Э.Р., Белоусов В.В., Агафонов Е.С. Инструментальная система автоматического поиска оптимальных программ поставок в системах послепродажного обслуживания изделий // Системы высокой доступности. М.: Радиотехника. 2013. Т 9. № 2. - С. 35-42.
8. Пугачев В.С., Синицын И.Н. Теория стохастических систем. М.: Изд-во «Логос», 2000 и 2003. (1 и 2 изд.)
9. Синицын И.Н. Канонические представления случайных функций и их применения в задачах компьютерной поддержки научных исследований. – М.: ТОРУС ПРЕСС, 2009.

УДК 004.03; +530.1
И.М. Гуревич, канд. техн. наук
Институт проблем информатики РАН, г. Москва, Россия,
ООО «ГЕТНЕТ Консалтинг», г. Москва, Россия
iggurevich@gmail.com

ИНФОРМАЦИОННЫЕ ОСНОВЫ И ОГРАНИЧЕНИЯ НАНОТЕХНОЛОГИЙ

Введение

«В 1959 г. крупнейший американский физик – лауреат Нобелевской премии Ричард Фейнман (Richard Feynman) заявил: «Пока мы вынуждены пользоваться атомарными структурами, которые предлагает нам природа». И добавил: «Но в принципе физик мог бы синтезировать любое вещество по заданной химической формуле». Знаменитая лекция Фейнмана, известная под названием «Там, внизу, ещё много места» считается сегодня стартовой точкой в борьбе за покорение наномира. В последние годы темпы научно-технического прогресса стали зависеть от использования искусственно созданных объектов нанометровых размеров 1 нанометр (нм) равен одной миллиардной доле метра или, что то же самое, одной миллионной доле миллиметра). Созданные на их основе вещества и объекты размером 1 -100 нм называют наноматериалами, а способы их производства и применения – нанотехнологиями. В самом широком смысле нанотехнологии – это исследования и разработки на атомном, молекулярном и макромолекулярном уровне в масштабе размеров от одного до ста нанометров; создание и использование искусственных структур, устройств и систем, которые в силу своих сверхмалых размеров обладают существенно новыми свойствами и функциями; манипулирование веществом на атомной шкале расстояний» [1].

«Конечная цель нанотехнологии заключается в производстве атомно-молекулярных конструкций с заданными

эксплуатационными свойствами и функциями. Авангардом нанотехнологии выступает механосинтез, представляющий собой технологию нанопроизводства, основанную на манипулировании отдельными атомами и молекулами. Очевидно, что успешное становление механосинтеза связано с созданием теории, обеспечивающей количественное описание связи свойств атомов и молекул с их структурой. Речь идет о создании теории, вооружающей практику методами и средствами, которые позволили бы осуществлять инженерные расчеты отдельных элементов наноконструкций и вместе с тем стать основой теоретического обеспечения нанопроектирования. В пределе такая теория должна дать ответ на вопрос какие исходные элементы (атомы и молекулы) надо выбрать для наносборки, каким образом и в какой последовательности их следует собирать для получения у создаваемого материала или изделия требуемых эксплуатационных свойств и функций. По сути, проблема целенаправленной наносборки – это проблема современного естествознания и в первую очередь проблема атомно-молекулярного строения вещества и межчастичных взаимодействий. Свойства вещества или искусственной атомно-молекулярной наноконструкции (механические, геометрические, электрические, магнитные, «химические») генетически обусловлены свойствами собственно атомов и молекул, используемых для ее построения. Располагая знаниями о строении атомов и механизмами образования атомно-молекулярных систем, можно приступить к разработке общей теории электронного строения, в том числе и стратегии наноконструирования» [2].

Взаимопроникновение физики и теории информации в ходе развития информатики сформировало новую синтетическую научную дисциплину «Физическая информатика», которая информационными методами исследует естественные (физические, химические и биологически) системы. «Физическая информатика» изучает информационные характеристики естественных систем, дает оценки

характеристик, формирует информационные ограничения на природные явления и процессы, определяет естественнонаучные следствия из законов информатики: физические законы и свойства природы. Данная научная дисциплина создана, в основном, в работах автора [3-76]. «Физическая информатика» также изучает информационные характеристики наносистем, дает оценки их характеристик, формирует информационные ограничения на нанотехнологии.

Информационные характеристики естественных систем

На произвольных естественных системах, объектах наряду с физическими характеристиками (массой, энергией, зарядом,...), определяются и информационные характеристики – информационная энтропия, информационная дивергенция, совместная информационная энтропия, информация связи, дифференциальная информационная емкость [3-7]. Информационная энтропия, информационная дивергенция определены на наблюдаемых и состояниях физической системы, совместная информационная энтропия определена на унитарных преобразованиях, связывающих состояния физической системы. Информация связи определена на наблюдаемых (наблюдаемой в квантовой механике называют любую физическую величину, которую можно измерить, причем результатами эксперимента обязательно должны являться действительные числа) и состояниях взаимодействующих подсистем физической системы и унитарных преобразованиях связывающих подсистемы.

Приведем, два примера, показывающие как на естественных системах определяется информационная энтропия. Согласно квантовой механике система, находящаяся в чистом состоянии, описывается волновой функцией или амплитудой вероятности, квадраты модулей которых определяют вероятность реализации ее состояний, а система, находящаяся в смешанном состоянии, описывается набо-

ром волновых функций или амплитуд вероятности, заданных с определенными вероятностями.

Дискретный случай. H - гильбертово пространство, $\psi \subset H$ - вектор пространства H. $\left|e_i\right\rangle$ i=1,…n - система базисных ортонормированных векторов пространства H. $\left\langle e_i \middle| e_j \right\rangle = \delta_{ij}$. $\delta_{ij} = \begin{pmatrix} 1, & \text{i=j} \\ 0, & \text{i}\neq\text{j} \end{pmatrix}$.

Физическая система отождествляется с вектором (волновой функцией) $\psi \subset H$: $\psi = \Sigma c_i \left|e_i\right\rangle$, $\sum\limits_i \left|c_i\right|^2 = 1$ [8].

Следовательно, на физической системе $\psi = \Sigma c_i \left|e_i\right\rangle$ определено вероятностное пространство: множество Ω элементарных событий (исходов) составляют базисные вектора $\left|e_i\right\rangle$; вероятностная мера задается квадратами модулей коэффициентов $c_i \Rightarrow p_i = \left|c_i\right|^2$ - вероятность реализации состояния e_i при измерении физической системы в базисе $\left(e_i\right)$. $\Sigma\ p_i = \Sigma\ \left|c_i\right|^2 = 1$. Информационная энтропия (по Шеннону) случайной величины x, принимающей значения x_i с вероятностями p_i ($\Sigma\ p_i = 1$) по определению равна

$$N = -\Sigma\ p_i \log_2 p_i \text{ бит.}$$

Информационная энтропия характеризует неопределенность содержащуюся в случайной величине x (до измерения) и/или информацию, содержащуюся в случайной величине x (после измерения). Таким образом, физи-

ческая система, отождествляемая с вектором (волновой функцией) $\psi = \Sigma c_i \left| e_i \right\rangle$, содержит

$$N = -\Sigma \; p_i \log_2 p_i = \Sigma \; \left|c_i\right|^2 \log_2 \left|c_i\right|^2 \text{ бит неопределенности (информации).}$$

Непрерывный случай. H - гильбертово пространство, $\psi(q) \subset H$ - вектор пространства H, здесь q - обобщенная координата. Физическая система отождествляется с вектором (волновой функцией) $\psi(q)$. Физическая система $\psi(q)$ наделена вероятностной структурой: множество Ω элементарных событий (исходов) составляют значения координаты q; вероятностная мера задается плотностью распределения (вероятности) $p(q) = \left|\psi(q)\right|^2$. $p(x_0)dx = P\left\{x_0 < x < x_0 + dx\right\}$ - вероятность реализации состояния q при измерении физической системы $\int p(q)dq = \int \left|\psi(q)\right|^2 dq = 1$. Информационная энтропия (по Шеннону) случайной величины q, характеризующейся плотностью распределения (вероятности) $p(q)$ $(\int p(q)dq = 1)$ по определению равна

$$N = -\int_0^a p(q) \log_2 p(q) dq \text{ бит.}$$

Информационная энтропия характеризует неопределенность содержащуюся в случайной величине q (до измерения) и/или информацию, содержащуюся в случайной величине q (после измерения). Таким образом, физическая система, отождествляемая с вектором (волновой функцией) $\psi(q)$, несет (содержит)

$N = -\int_0^a p(q)\log_2 p(q)dq$ бит неопределенности (информации).

Связь информации с энтропией, энергией, массой. Физическая энтропия S пропорциональна логарифму числа микросостояний, в которых может находиться система $S = k \ln P$. $k = 1{,}38\ 10^{-16}$ эрг/К = $1{,}38\ 10^{-23}$ Дж/К – постоянная Больцмана; а P – статистический вес (число микросостояний). Если система имеет 2 состояния, то энтропия системы равна $S = k \ln 2 = 1{,}38 \cdot 10^{-16} \ln 2$ эрг/К = $1{,}38 \cdot 10^{-23} \ln 2$ Дж/К. Объем информации в системе из двух равновероятных состояний в информационных единицах равен одному биту, а в единицах энтропии равен постоянной Больцмана, умноженной на $\ln 2$. В общем случае энтропия системы, содержащей I бит информации, равна постоянной Больцмана, умноженной на $I \ln 2$: $S = k I \ln 2$.

Энергия, требуемая для передачи, чтения, записи одного бита при температуре T, не может быть меньше величины $E_{min} = kT \ln 2$. Соответственно, учитывая формулу Эйнштейна, масса, требуемая для передачи, чтения или записи одного бита при температуре T, не может быть меньше величины $M_{min} = (kT \ln 2)/c^2$.

Информационные характеристики естественных систем (например: объем информации в расширяющейся Вселенной, квадратичная зависимость информации в черной дыре от массы) исследуются с привлечением информационных методов. При этом используются известные, общепризнанные физические модели. Так как информационные характеристики естественных систем связаны с физическими характеристиками, то при исследовании информационных характеристик одновременно исследуются физические характеристики (например: структура, излучение черной дыры, масса начальных неоднородностей Вселенной).

Информационные законы природы (законы информатики) – основы исследования естественных систем

Физические исследования проводятся около трех тысяч лет и для описания разнообразных природных явлений и процессов накоплен значительный объем физических понятий и методов.

По мнению автора, существуют законы более общие, чем физические – законы информатики [3-7], *определяющие, ограничивающие физические явления, процессы и технологии, законы, предшествующие физическим законам.*

Законы информатики

Закон простоты сложных систем

Реализуется, выживает, отбирается тот вариант сложной системы, который обладает наименьшей сложностью.

Закон простоты сложных систем реализуется природой в ряде конструктивных принципов:

- «бритва Оккама»;
- иерархического модульного построения сложных систем;
- симметрии;
- симморфоза (равнопрочности, однородности);
- устойчивости;
- полевого взаимодействия (взаимодействия через носитель или взаимодействия через состояние пространства-времени, например, кривизну пространства-времени);
- экстремальной неопределенности (функции распределения характеристик, параметров, имеющих неопределенные значения, имеют экстремальную неопределенность).

Важной реализацией закона простоты сложных систем является:

Закон сохранения неопределенности (информации)

Неопределенность (информация) изолированной (замкнутой) системы сохраняется при физически реализуемых преобразованиях и только при физически реализуемых преобразованиях.

Закон конечности информационных характеристик сложных систем

Все виды взаимодействия между системами, их частями и элементами имеют конечную скорость распространения. Ограничена также скорость изменения состояний элементов системы.

- В любой системе координат информация о событии всегда конечна.
- Длительность сигнала ΔT всегда больше нуля ($\Delta T > 0$).
- Информация о координатах физических систем в нашем мире ограничена 333 битами.

Закон необходимого разнообразия Эшби

Для эффективного функционирования системы разнообразие управляющего органа должно быть не менее разнообразия объекта управления.

Отметим, что неопределенность (информация) является основной характеристикой разнообразия системы.

Закон необходимого разнообразия Эшби также реализуется в ряде конкретных принципов:

- теоремы Шеннона,
- теорема Котельникова,
- теорема Холево,
- теорема Брюллиена,
- теорема Марголиса–Левитина.

Теорема Геделя о неполноте

В достаточно богатых теориях (включающих арифметику) всегда существуют недоказуемые истинные

выражения.

Характер изменения, эволюции сложных систем определяют:

Закон роста сложности систем

В ходе эволюции системы ее неопределенность (информация в ней) растет.

Закон Онсагера максимизации убывания энтропии

Если число всевозможных форм реализации процесса, не единственно, то реализуется та форма, при которой энтропия системы растет наиболее медленно. Иначе говоря, реализуется та форма, при которой максимизируется убывание энтропии или рост информации, содержащейся в системе.

Принцип ле Шателье

Внешнее воздействие, выводящее систему из равновесия, вызывает в ней процессы, стремящиеся ослабить результаты этого воздействия.

Методика иерархической оценки объема информации в естественных объектах

Методика оценки объема информации в естественных объектах иерархической структуры заключается в следующем [5-8]. Сначала оценивается объем информации в объектах нижнего уровня (лептонах и кварках). Согласно основному принципу квантовой механики Цайлингера считаем, что в объектах нижнего уровня – фундаментальных частицах содержится 1 бит информации. Далее оценивается объем неопределенности (информации) в объектах второго уровня, который равен сумме объемов информации объектов нижнего уровня плюс объем информации, заключенной в структуре объекта второго уровня иерархии (мезоны, барионы). Объем информации в структуре объекта второго уровня оценивается по волновой функции объекта второго уровня и/или по графу, отображающему его

структуру. Затем оценивается объем информации в объектах следующего уровня, который равен сумме объемов информации, входящих в его состав объектов предыдущих уровней, плюс объем информации, заключенной в структуре объекта последующего уровня иерархии (атомы). Объем информации в структуре объекта третьего уровня оценивается по волновой функции объекта третьего уровня. И так далее: молекулы, твердые тела, звезды, галактики, Вселенная. Аминокислоты, азотистые основания, клетки, организмы, популяции, общества, цивилизации. В ряде случаев, необходимо учитывать пространственную неопределенность (информацию).

Оценка объема информации в структуре молекул

Предлагается использовать оценку объема информации в структуре графа, соответствующей структуре молекулы $I_{gr} = \sum_{i=1}^{m} n_i \log_2 \nu_i$ приведенную в работах [5-8]. Здесь где m - количество классов топологически эквивалентных вершин графа (вершин заданной степени ν_i), n_i – число вершин графа степени ν_i. Данная оценка является оценкой сверху. Рассматриваемое количество направлений выхода из каждой вершины обеспечивает обход графа с использованием всех возможных путей (направлений перехода) от атома к атому.

Оценка объема информации в структуре графа, соответствующей структуре молекулы, использующая степени вершин, уменьшенные на единицу,

$$I'_{Gr} = \sum n_j \log_2(\nu_j - 1),$$

является оценкой снизу.

Оценка объема информации в структурах, формируемых атомами углерода (на один атом)

В алмазе, углеводородах – $\log_2 4 = 2$ бита,

в графе́не , фуллеренах, углеродных нанотрубках – $\log_2 3 = 1{,}5849$ бит,

в графите – $\log_2 3 = 1{,}5849$ бит.

Для сравнения объем информации в атоме водорода равен 12, 422 бита, а объем информации в структуре молекулы воды равен 1 биту. Объем информации в структуре молекулы метана равен 2 бита, в структуре молекулы бутана равен 8 бит.

Оценки объема информации в атомах приведены в таблице 1. Здесь A - заряд (атомный номер) - число протонов, электронов; B - число нейтронов; C - число протонов и нейтронов; D - объем информации в протоне, нейтроне; E - объем информации в электроне; F - чсло электронов; G - объем информации в частицах, составляющих атом; H - число орбиталей; I - объем информации в одной орбитали; J - объем информации в структуре атома; K - объем информации в спине электрона; L - объем инфор мации в атоме).

Таблица 1 - Оценки объема информации в атомах

A	B	C	D	E	F	G	H	I	J	K	L
H 1	0	1	9,422	1	1	10,422	1	1	1	1	12,422
He 2	2	4	9,422	1	2	39,688	1	1	1	1	42,688
Li 3	3	6	9,422	1	3	59,532	2	1	2	1	64,532
Be 4	5	9	9,422	1	4	88,798	2	1	2	1	94,798
B 5	5	10	9,422	1	5	99,22	3	1	3	1	107,22
C 6	5	11	9,422	1	6	109,642	3	1	3	1	118,642
N 7	7	14	9,422	1	7	138,908	4	1	4	1	149,908
O 8	7	15	9,422	1	8	149,33	4	1	4	1	161,33

A	B	C	D	E	F	G	H	I	J	K	L
F 9	8	17	9,422	1	9	169,174	5	1	5	1	183,174
Ne 10	10	20	9,422	1	10	198,44	5	1	5	1	213,44
Na 11	11	22	9,422	1	11	218,284	6	1	6	1	235,284
Mg 12	12	24	9,422	1	12	238,128	6	1	6	1	256,128
Al 13	13	26	9,422	1	13	257,972	7	1	7	1	277,972
Si 14	14	28	9,422	1	14	277,816	7	1	7	1	298,816
P 15	15	30	9,422	1	15	297,66	8	1	8	1	320,66
S 16	16	32	9,422	1	16	317,504	8	1	8	1	341,504
Cl 17	18	35	9,422	1	17	346,77	9	1	9	1	372,77
Ar 18	21	39	9,422	1	18	385,458	9	1	9	1	412,458
K 19	20	39	9,422	1	19	386,458	10	1	10	1	415,458
Ca 20	20	40	9,422	1	20	396,88	10	1	10	1	426,88
Sc 21	23	44	9,422	1	21	435,568	11	1	11	1	467,568
Ti 22	25	47	9,422	1	22	464,834	11	1	11	1	497,834
V 23	27	50	9,422	1	23	494,1	12	1	12	1	529,1
Cr 24	27	51	9,422	1	24	504,522	12	1	12	1	540,522
Mn 25	29	54	9,422	1	25	533,788	13	1	13	1	571,788
Fe 26	29	55	9,422	1	26	544,21	13	1	13	1	583,21
Co 27	31	58	9,422	1	27	573,476	14	1	14	1	614,476
Ni 28	30	58	9,422	1	28	574,476	14	1	14	1	616,476
Cu 29	34	63	9,422	1	29	622,586	15	1	15	1	666,586
Zn 30	35	65	9,422	1	30	642,43	15	1	15	1	687,43
Ga 31	38	69	9,422	1	31	681,118	16	1	16	1	728,118
Ge 32	40	72	9,422	1	32	710,384	16	1	16	1	758,384
As 33	41	74	9,422	1	33	730,228	17	1	17	1	780,228
Se 34	44	78	9,422	1	34	768,916	17	1	17	1	819,916

A	B	C	D	E	F	G	H	I	J	K	L
Br 35	44	79	9,422	1	35	779,338	18	1	18	1	832,338
Kr 36	47	83	9,422	1	36	818,026	18	1	18	1	872,026
Rb 37	48	85	9,422	1	37	837,87	19	1	19	1	893,87
Sr 38	49	87	9,422	1	38	857,714	19	1	19	1	914,714
Y 39	49	88	9,422	1	39	868,136	20	1	20	1	927,136
Zr 40	51	91	9,422	1	40	897,402	20	1	20	1	957,402
Nb 41	51	92	9,422	1	41	907,824	21	1	21	1	969,824
Mo 42	53	95	9,422	1	42	937,09	21	1	21	1	1000,09
Tc 43	54	97	9,422	1	43	956,934	22	1	22	1	1021,93
Ru 44	57	101	9,422	1	44	995,622	22	1	22	1	1061,62
Rh 45	57	102	9,422	1	45	1006,04	23	1	23	1	1074,04
Pd 46	60	106	9,422	1	46	1044,73	23	1	23	1	1113,73
Ag 47	60	107	9,422	1	47	1055,15	24	1	24	1	1126,15
Cd 48	64	112	9,422	1	48	1103,26	24	1	24	1	1175,26
In 49	65	114	9,422	1	49	1123,11	25	1	25	1	1197,11
Sn 50	68	118	9,422	1	50	1161,8	25	1	25	1	1236,8
Sb 51	70	121	9,422	1	51	1191,06	26	1	26	1	1268,06
Te 52	75	127	9,422	1	52	1248,59	26	1	26	1	1326,59
I 53	73	126	9,422	1	53	1240,17	27	1	27	1	1320,17
Xe 54	77	131	9,422	1	54	1288,28	27	1	27	1	1369,28
Cs 55	77	132	9,422	1	55	1298,7	28	1	28	1	1381,7
Ba 56	81	137	9,422	1	56	1346,81	28	1	28	1	1430,81
La* 57	81	138	9,422	1	57	1357,24	29	1	29	1	1443,24
Ce 58	82	140	9,422	1	58	1377,08	29	1	29	1	1464,08
Pr 59	81	140	9,422	1	59	1378,08	30	1	30	1	1467,08
Nd 60	84	144	9,422	1	60	1416,77	30	1	30	1	1506,77

A	B	C	D	E	F	G	H	I	J	K	L
Pm 61	86	147	9,422	1	61	1446,03	31	1	31	1	1538,03
Sm 62	88	150	9,422	1	62	1475,3	31	1	31	1	1568,3
Eu 63	88	151	9,422	1	63	1485,72	32	1	32	1	1580,72
Gd 64	93	157	9,422	1	64	1543,25	32	1	32	1	1639,25
Tb 65	93	158	9,422	1	65	1553,68	33	1	33	1	1651,68
Dy 66	96	162	9,422	1	66	1592,36	33	1	33	1	1691,36
Ho 67	97	164	9,422	1	67	1612,21	34	1	34	1	1713,21
Er 68	99	167	9,422	1	68	1641,47	34	1	34	1	1743,47
Tm 69	99	168	9,422	1	69	1651,9	35	1	35	1	1755,9
Yb 70	103	173	9,422	1	70	1700,01	35	1	35	1	1805,01
Lu 71	103	174	9,422	1	71	1710,43	36	1	36	1	1817,43
Hf 72	106	178	9,422	1	72	1749,12	36	1	36	1	1857,12
Ta 73	107	180	9,422	1	73	1768,96	37	1	37	1	1878,96
W 74	109	183	9,422	1	74	1798,23	37	1	37	1	1909,23
Re 75	111	186	9,422	1	75	1827,49	38	1	38	1	1940,49
Os 76	114	190	9,422	1	76	1866,18	38	1	38	1	1980,18
Ir 77	115	192	9,422	1	77	1886,02	39	1	39	1	2002,02
Pt 78	117	195	9,422	1	78	1915,29	39	1	39	1	2032,29
Au 79	117	196	9,422	1	79	1925,71	40	1	40	1	2044,71
Hg 80	120	200	9,422	1	80	1964,4	40	1	40	1	2084,4
Ti 81	123	204	9,422	1	81	2003,09	41	1	41	1	2125,09
Pb 82	125	207	9,422	1	82	2032,35	41	1	41	1	2155,35
Bi 83	125	208	9,422	1	83	2042,78	42	1	42	1	2167,78
Po 84	125	209	9,422	1	84	2053,2	42	1	42	1	2179,2
At 85	137	222	9,422	1	85	2176,68	43	1	43	1	2304,68
Rn 86	145	231	9,422	1	86	2262,48	43	1	43	1	2391,48

A	B	C	D	E	F	G	H	I	J	K	L
Fr 87	136	223	9,422	1	87	2188,11	44	1	44	1	2319,11
Ra 88	138	226	9,422	1	88	2217,37	44	1	44	1	2349,37
Ac 89	138	227	9,422	1	89	2227,79	45	1	45	1	2361,79
Th 90	142	232	9,422	1	90	2275,9	45	1	45	1	2410,9
Pa 91	140	231	9,422	1	91	2267,48	46	1	46	1	2404,48
U 92	146	238	9,422	1	92	2334,44	46	1	46	1	2472,44
Np 93	144	237	9,422	1	93	2326,01	47	1	47	1	2466,01
Pu 94	150	244	9,422	1	94	2392,97	47	1	47	1	2533,97
Am 95	148	243	9,422	1	95	2384,55	48	1	48	1	2527,55
Cm 96	151	247	9,422	1	96	2423,23	48	1	48	1	2567,23
Bk 97	150	247	9,422	1	97	2424,23	49	1	49	1	2570,23
Cf 98	153	251	9,422	1	98	2462,92	49	1	49	1	2609,92
Es 99	153	252	9,422	1	99	2473,34	50	1	50	1	2622,34
Fm 100	157	257	9,422	1	100	2521,45	50	1	50	1	2671,45
Md 101	157	258	9,422	1	101	2531,88	51	1	51	1	2683,88
No 102	157	259	9,422	1	102	2542,3	51	1	51	1	2695,3
Lr 103	157	260	9,422	1	103	2552,72	52	1	52	1	2707,72
Rf 104	157	261	9,422	1	104	2563,14	52	1	52	1	2719,14
Db 105	157	262	9,422	1	105	2573,56	53	1	53	1	2731,56
Sg106	157	263	9,422	1	106	2583,99	53	1	53	1	2742,99
Bh 107	157	264	9,422	1	107	2594,41	54	1	54	1	2755,41
Hs 108	157	265	9,422	1	108	2604,83	54	1	54	1	2766,83
Mt 109	157	266	9,422	1	109	2615,25	55	1	55	1	2779,25
Uun110	161	271	9,422	1	110	2663,36	55	1	55	1	2828,36

Заметим, что в среднем в атомах на 1 бит информации используется $\approx 1{,}69 \cdot 10^{-28}$ кг массы вещества (например: в атоме водорода $\approx 1{,}6 \cdot 10^{-28}$кг, в атоме лития $\approx 1{,}93 \cdot 10^{-28}$кг, а средне-квадратичное отклонение равно $\approx 7{,}86 \cdot 10^{-30}$, что не превосходит 5% от среднего значения). Это примерно в $\approx 6 \cdot 10^{11}$ раз больше минимальной массы, необходимой для формирования 1 бита микроинформации при температуре 2,7K, равной $m_{\min} = \frac{kT}{c^2} \ln 2 = \frac{2{,}7k}{c^2} \ln 2 \approx 10^{-40}$ кг.

Затраты массы на 1 бит информации в структуре молекулы метана больше затрат массы на 1 бит информации в атоме водорода в 205,596 раз. Затраты массы на 1 бит информации в структуре молекулы воды больше затрат массы на 1 бит информации в структуре молекулы метана в 2,25 раз. Затраты массы на 1 бит информации в структуре молекулы воды больше затрат массы на 1 бит информации в атоме водорода в 91,4 раз. Таким образом, затраты массы на 1 бит информации в структуре молекул на два порядка больше затрат массы на 1 бит информации в атомах.

Свойства и возможности атомно-молекулярных наноконструкций определяются, главным образом, вышеприведенными оценками объема информации в атомах, используемых для их построения.

Информационные мировые константы. Информационные ограничения на формирование и развитие естественных и искусственных объектов

Физические мировые константы [10]: постоянная Планка h, гравитационная постоянная G, скорость света c, постоянная Больцмана k, … - определяют характеристики и свойства Вселенной и всех физических систем, их формирование, развитие, взаимодействие и взаимопревра-

щения, ... Они также определили возможность вения жизни.

Развитие производства запоминающих устройств и микрочипов связано с увеличением плотности записи информации на эти носители. Может создаться впечатление, что изобретение новых технологий увеличения плотности записи процесс постоянный.

«Человек всегда стремился к большему, желая не просто повторить изобретения природы но и превзойти их. До сих пор это ему не удавалось и, лишь с освоением нанотехнологий он может получить реальные шансы на воплощение своей давней «бредовой» мечты – присвоение функции Творца Вселенной, связанной с возможностью по своей воле создавать новый мир на основе биоорганомики, соединившей физику и молекулярную биологию» [11].

Однако, существуют фундаментальные ограничения, которые неминуемо остановят этот процесс и поставят земную цивилизацию перед выбором: либо прекратить запоминать новую информацию, либо «стирать» накопленную ранее. Экспоненциальный рост объемов хранимой на нашей планете информации может привести к такому моменту неожиданно скоро. Возможности технологии ограничены конструкцией и свойствами атомов и молекул и информационными константами. В работах автора [3-8] показано, что объем информации во Вселенной ограничен, объем информации в атомах, аминокислотах и азотистых основаниях, их дифференциальная информационная емкость, масса атома водорода, структура и разность энергий его базисных состояний накладывают ограничения на память и быстродействие вычислительных устройств, информационных систем и затраты, необходимые для записи и передачи информации. Информационные ограничения определяют процессы формирования и развития естественных и искусственных объектов, в том числе, срок действия закона Мура и его аналогов.

Фундаментальные ограничения на объем информации в естественных системах. Впервые оценка объема информации в нашей Вселенной 10^{90} бит была дана в 1989 году в работе автора [3]. Данное ограничение является фундаментальным ограничением на сложность естественных и искусственных систем и является информационной мировой константой – *информационной границей* IB = 10^{90} бит. Эта же оценка была получена Lloyd Seth в 2001 году [12]. В работах [4-8] представены оценки объема информации в элементарных частицах, атомах, молекулах, звездах, галактиках, …и сделаны следующие выводы:

Причина и источник формирования информации – расширение Вселенной и ее исходная неоднородность. При расширении Вселенной изменяется ее фазовое состояние (симметрия) и кривизна пространства; формируются различные типы неоднородностей массы и энергии, в частности, возникают фундаментальные и элементарные частицы, галактические, звездные, планетные системы; формируются аминокислоты, азотистые основания, белки, ДНК, организмы, цивилизации. Наиболее подходящими для формирования и хранения информации структурными единицами материи являются фермионы, а для передачи информации – бозоны. При нарушении симметрии между слабым и электромагнитным взаимодействиями во Вселенной формируется 10^{90} бит информации. «Информационный» механизм формирования частиц в инфляционной Вселенной порождает количество частиц, сравнимое с общепринятой оценкой числа частиц во Вселенной, порядка 10^{80}–10^{90}. Степенное расширение Вселенной порождает неоднородности в объеме $\approx 10^{99} - 10^{107}$ бит, из них в обычном веществе $\approx 10^{90}$ бит. Во Вселенной, в звездах содержится около 10^{80} бит информации. Если в ядрах галактик находятся черные дыры массой $\approx 10^{6} - 10^{10}$ солнечных масс, то объем информации во Вселенной составляет $\approx 10^{99} - 10^{107}$ бит. Минимально возможный объем информации во Вселенной

с преобладанием вещества $\approx 1{,}7 \cdot 10^{79}$ бит, а с преобладанием излучения $\approx 10^{91}$ бит. Максимально возможный объем информации во Вселенной $\approx 10^{120}$ бит. Рост объема информации при степенном расширении Вселенной $\propto \log_2 t$. Уменьшение плотности информации при степенном расширении Вселенной $\propto \frac{\log_2 t}{t^2}$. Рост объема информации при экспоненциальном расширении Вселенной $\propto \alpha t$. Уменьшение плотности информации при экспоненциальном расширении Вселенной $\propto \frac{t}{e^{3\alpha t}}$. Панов А.Д. в статье «Квантовая реальность, вычисления и искусственный интеллект» (в настоящем сборнике) использует информационную мировую константу – *информационную границу* IB = 10^{90} бит в качестве ограничения на возможности искусственного интеллекта. Он отмечает: «Фактически, для вычисления поведения некоторых практически важных квантовых систем требуются такие мощности классического компьютера или такие объемы вычислений, которые нельзя реализовать не только практически, но и принципиально, так как для вычислений будет необходимо время, превышающее возраст Вселенной, или размер компьютера будет таков, что его нельзя будет уместить внутри космологического горизонта событий. Например, для того, чтобы с использованием квантового алгоритма Шора разложить на простые множители 1000-значное двоичное число (обычная задача для перспективных кантовых компьютеров), квантовому компьютеру требуется память всего в несколько тысяч квантовых ячеек памяти – кубитов, в то время как классическому компьютеру для представления состояния такого квантового компьютера потребуется память порядка 21000 комплексных чисел – а это уже на много порядков больше объема информации, которая может быть записана во всем обычном веществе видимой части Вселенной (10^{90} бит).

Поэтому реальные классические компьютерные симуляторы квантовых вычислений могут работать только с очень малоразмерными системами».

Фундаментальные ограничения на информационную емкость естественных и искусственных систем. Существует несколько типов материи с разной зависимостью объема информации (информационной емкости) от массы. Для обычного вещества $I = \beta M$, $f'(M) = \beta$. Дифференциальная информационная емкость обычного вещества не зависит от его массы. В настоящее время системы строятся из обычного вещества. Для обычного вещества, содержащего в атоме, молекуле $I_{at,\,mol}$ бит информации на 1 бит информации необходима масса $\beta = \dfrac{I_{at,mol}}{m_{at,mol}}$. Поэтому $I = \dfrac{I_{at,mol}}{m_{at,mol}} M$, $f'(M) = \dfrac{I_{at,mol}}{m_{at,mol}}$. Таким образо, дифференциальная информационная емкость атомов разных элементов и, соответственно молекул, примерно одинакова.

Дифференциальная информационная емкость атомов, аминокислот, азотистых оснований определяют фундаментальные ограничения на информационную емкость устройств хранения данных [13-14].

Ограничение, накладываемое дифференциальной информационной емкостью неживой материи. Атомы - простейшие средства хранения информации. На основе оценки дифференциальной информационной емкости атомов – $\approx 10^{-28}$ бит/кг определяется нижняя граница G дифференциальной информационной емкости V искусственных устройств хранения данных $G \approx 10^{-28}$ бит/кг. Следовательно, информационная емкость I устройств хранения данных, построенных на базе атомов, не превосходит

$10^{28} M$ бит, где M - масса устройства хранения данных. Т.к. в настоящее время дифференциальная информационная емкость устройств хранения данных $\approx 10^{14}$ бит/кг, то для устройств, построенных на базе атомов, она может быть повышена не более, чем в $\approx 10^{14}$ раз. Ограничение (10^{28} бит/кг) является фундаментальным ограничением, накладываемым природой на информационную емкость естественных и искусственных систем и определяется как информационная мировая константа – *постоянная памяти* IM = 10^{28} бит/кг.

Ограничение, накладываемое характеристиками атома водорода. На один бит в атоме водорода, рассматриваемого как q-бит, природа тратит $m_H = 1,67 \cdot 10^{-27}$ кг. Следовательно, информационная емкость I устройств хранения данных, построенных на базе атомов водорода, используемых как q-биты, не превосходит $6 \cdot 10^{26} M$ бит, где M - масса устройства хранения данных масса водорода.

Ограничение, накладываемое дифференциальной информационной емкостью живой материи. Белки, ДНК – простейшие естественные средства хранения информации. На основе оценки дифференциальной информационной емкости белков, ДНК – $\approx 10^{-25}$ кг/бит определяется нижняя граница IM дифференциальной информационной емкости устройств хранения данных – и устройств построенных на базе комбинаций атомов, $IM \approx 10^{25}$ бит/кг. Следовательно, информационная емкость устройств хранения данных, построенных на базе комбинаций атомов, не превосходит $10^{25} M$ бит, где M - масса устройства хранения данных. Т.к. в настоящее время дифференциальная информационная емкость устройств хранения данных

$\approx 10^{14}$ бит/кг, то для устройств, построенных на базе комбинаций атомов, она может быть повышена не более, чем в $\approx 10^{11}$ раз.

Фундаментальные ограничения на производительность естественных и искусственных систем. Разность энергий базисных состояний атома водорода, рассматриваемого как q-бит, накладывают фундаментальные ограничения на быстродействие вычислительных устройств. Согласно теореме Н. Марголиса и Л. Левитина [15] общее количество элементарных действий, которые система может выполнить в секунду, ограничено энергией: $k_{op/s} = 2E / \hbar$, где E – превышение средней энергии системы над энергией нижнего состояния или энергия активации $\hbar = \frac{h}{2\pi}$ = 1,0545 10^{-34} с Дж – уменьшенная постоянная Планка. Число операций, выполняемых атомом водорода как q-битом, ограничено $k_{op/s} = 2\Delta E / \hbar \approx 1,5 \cdot 10^{12}$ операций в секунду, а производительность компьютера массой один килограмм, построенного из атомов водорода, не превышает 10^{39} (оп/с)/кг.

Данное ограничение является наиболее сильным фундаментальным ограничением, на быстродействие вычислений естественных и искусственных систем и опредеяется как *постоянная быстродействия* IS = 10^{39} (оп/с)/кг.

Фундаментальные ограничения на характеристики компьютеров. Память компьютера массой один килограмм, построенного из атомов водорода, не превышает 0,6 10^{27} бит, а его производительность не превышает 10^{39} оп/с.

В 1986 г. американский инженер Эрик Дрекслер предложил использовать для производства наноустройств

механические машины соответствующих (100-200 нм) размеров – нанороботы [11, 16-17]. Эти роботы должны были собирать устройство непосредственно из атомов, и были названы ассемблерами – сборщиками. Сборщик оснащен манипуляторами длиной в несколько десятков нанометров, двигателем для перемещения манипуляторов и робота, а также автономным источником энергии. Наноробот должен состоять из нескольких десятков тысяч деталей, а каждая деталь – из одной-двух сотен атомов. Важнейший узел наноробота – бортовой компьютер, который управляет работой всех механизмов, определяя какой атом или какую молекулу следовало захватить манипулятором и в какое место будущего устройства их поставить. Линейные размеры этого компьютера не должны были превышать 40-50 нм, а его объем равен $V_C \approx 10^{-22}$ м^3. В компьютере можно разместить 10^8 атомов водорода, следовательно, память компьютера наноробота, построенного из атомов водорода, не превышает 10^8 бит, а его производительность $\leq 10^{20}$ оп/с.

Фундаментальные ограничения на энергию хранения и передачи информации. «К фундаментальным понятиям теории информации относится термодинамический предел для энергии переключения классического логического элемента $\left(P \cdot \tau\right)_{\min}$, определяемый как предельное значение минимальной работы $W_{\min} = \Delta F$, которую необходимо сообщить логическому элементу для того, чтобы термодинамически обратимым образом перевести его в состояние, отличающегося от исходного только на один бит информационной энтропии.

$$\left(P \cdot \tau\right)_{\min} = \Delta F_{\min} = kT \cdot \ln 2 \approx 3 \cdot 10^{-21}$$ Дж/бит.

Информационный процесс в логическом элементе можно охарактеризовать также энергией, называемой энергетической ценой одного бита. Для случая наличия шумов предельное значение энергетической цены равно $kT \ln 2 \approx 3 \cdot 10^{-21}$ Дж/бит. $k = 1,3806489 \cdot 10^{-23}$ Дж/К.

Поскольку $E = mc^2$, то к фундаментальным понятиям теории информации следует также отнести термодинамический предел для массы переключения классического логического элемента $M_{\min}$, определяемый как предельное значение минимальной массы $M_{\min} = \Delta m$, которую необходимо передать логическому элементу для того, чтобы термодинамически обратимым образом перевести его в состояние, отличающегося от исходного только на один бит информационной энтропии.

В единицах массы термодинамический предел равен

$$M_{\min} = \frac{E_{\min}}{c^2} = \frac{\left(P \cdot \tau\right)_{\min}}{c^2} = \frac{k \cdot \ln 2}{c^2} T \text{ кг/бит.}$$

При T = 300K $M_{\min} = \frac{k \ln 2}{c^2} \cdot \mathrm{T} \approx 3 \cdot 10^{-38}$ кг/бит.

Данное ограничение является наиболее сильным фундаментальным ограничением, накладываемым природой на затраты на запись и передачу информации в естественных и искусственных системах и является информационной мировой константой – постоянной *информационных затрат* IC = $3 \cdot 10^{-38}$ кг/бит.

Зако́н Му́ра – эмпирическое наблюдение, изначально сделанное Гордоном Муром, согласно которому (в современной формулировке) количество транзисторов, размещаемых на кристалле интегральной схемы, удваивается каждые 24 месяца. Часто цитируемый интервал в 18 месяцев связан с прогнозом Давида Хауса из Intel, по мнению которого *производительность* процессоров должна удваиваться каждые 18 месяцев из-за сочетания роста количества транзисторов и быстродействия каждого из них. В 2007 году Мур заявил, что закон, очевидно, скоро перестанет действовать из-за атомарной природы вещества и ограничения скорости света.

Поскольку информационная емкость I искусственных устройств хранения данных, построенных на базе атомов, не превосходит $10^{28} M$ бит, где M - масса устройства хранения данных, непосредственно используемая для хранения. Т.к. в настоящее время дифференциальная информационная емкость устройств хранения данных $\approx 10^{14}$ бит/кг, то для устройств, построенных на базе атомов, она может быть повышена не более, чем в $\approx 10^{14}$ раз. Для оценки принимаем, что емкость памяти удваивается каждые 24 месяца, каждые два года. Тогда предельное значение емкости памяти $10^{28} M$ бит, будет достигнута через 46,5 лет.

Заключение

Таким образом, наряду с физическими мировыми константами существуют информационные мировые константы:

- информационная граница (information border, limit) IB = 10^{90} бит,
- постоянная памяти (constant memory) M = 10^{28} бит/кг,
- постоянная быстродействия (constant speed, performance) IS = 10^{39} (оп/с)/кг,
- постоянная информационных затрат (information costs) IC= $3 \cdot 10^{-38}$ кг/бит.

Данные константы следует добавить в ряд мировых констант: постоянная Планка h , гравитационная постоянная G , скорость света c , постоянная Больцмана k , …

Информационные мировые константы определяют процессы формирования и развития естественных и искусственных объектов, в том числе, определяют перспективы развития наномехатроники.

Достичь данных мировых констант при исследовании и создании информационных систем невозможно – это ограничения, которые необходимо учитывать, это пределы к которым необходимо стремиться.

Список литературы

1. Евгений Биргер Примерная краткая история нанонауки и нанотехнологии.
http://www.washprofile.org/
2. Потапов А.А. Научные основы нанотехнологии // Институт динамики систем и теории управления СО РАН: http://85.142.23.144/packages/mifi/74A4B99F-5F63-4C8E-83F9-88E465A1DBD0/1.0.0.0/file.pdf
3. Гуревич И.М. Законы информатики – основа исследований и проектирования сложных систем связи и управления. Метод. пособие. - М.: ЦООНТИ «Экос», 1989. - 60 с.
4. Гуревич И.М. «Законы информатики – основа строения и познания сложных систем». Издание второе уточненное и дополненное. - М.: «Торус Пресс», 2007. - 400 с.
5. Гуревич И.М. Информационные характеристики физических систем. Издание второе уточненное и дополненное. – Севастополь: «Кипарис». Севастополь. 2010. 260 с.
6. Гуревич И.М., Урсул А.Д. Информация - всеобщее свойство материи: характеристики, оценки, ограничения, следствия. - М.: ЛИБРОКОМ, 2011. - 312 с.
7. Гуревич Игорь. ФИЗИЧЕСКАЯ ИНФОРМАТИКА. LAP Lambert Academic Publishing. (26.01.2012). 288с. ISBN: 978-3-8473-3873-4.
8. Igor Gurevich. Some works on physical informatics. LAP Lambert Academic Publishing. 2012-12-04 - ISBN-13: 978-3-659-30498-9.
9. Ландау Л.Д., Лившиц Е.М. Квантовая механика. Нерелятивистская теория. - М.: Наука, 1974.
10. Физические константы. Физический энциклопедический словарь. Т. 5. – М.: Советская энциклопедия, 1966. - С. 315-316.
11. Теряев Е.Д., Филимонов Н.Б. Наномехатроника: состояние, проблемы, перспективы // Мехатроника, автоматизация, управление. 2010. № 1. - С. 2-14.
12. Lloyd Seth. Computational capacity of the universe. arXiv:quant-ph/0110141 v1 24 Oct 2001.

13. Гуревич И.М. Фундаментальные ограничения на информационные характеристики систем // 3-я мультиконф. по проблемам управления. 7-я науч.-техн. конф. «Мехатроника, автоматизация, управление» (МАУ-2010). - Санкт-Петербур, 2010. - С. 144-147.
14. Гуревич И.М. Фундаментальные ограничения на емкость устройств хранения данных // Системы и средства информатики. Вып. 20, № 2. М.: ИПИ РАН, 2010. - С. 240-254.
15. Margolus N., Levitin L. Phys. Comp. 96. T. Toffoli, M. Biafore, J. Leao, eds. (NECSI, Boston) 1996; Physica D 120, 188-195 (1998).
16. Дрекслер К.Э. Машины создания. Грядущая эра нанотехнологии:
http://filosof.historic.ru/books/item/f00/s00/z0000328
17. Эрлих Г. В. Мифы нанотехнологий. 4 июня 2010:
http://www.nanometer.ru/2010/06/04/12756380321857_214259.html
18. Валиев К.А., Кокин А.А. Квантовые компьютеры: Надежда и реальность. Научно-издательский центр «Регулярная и хаотическая динамика». - Москва-Ижевск: 2001.
19. Валиев К.А. Квантовые компьютеры и квантовые вычисления // УФН. Т. 175. № 1, 2005.
20. Gordon E. Moore. Cramming more components onto integrated circuits. Electronics, Volume 38, Number 8, April 19, 1965.
21. Закон Мура
http://ru.wikipedia.org/wiki/%D0%97%D0%B0%D0%BA%D0%BE%_D0%BD_%D0%9C%D1%83%D1%80%D0%B0

УДК 530.1
А.Л. Леонтович, канд. физ.-мат. наук, доц.
М.П. Евстигнеев, д-р физ.-мат. наук., проф.
И.Ш. Шенгелая, канд. филос. наук, доц.
Севастопольский национальный технический университет, г. Севастополь, Украина

НЕКОТОРЫЕ ФИЛОСОФСКИЕ, МАТЕМАТИЧЕСКИЕ И ФИЗИЧЕСКИЕ АСПЕКТЫ ПРИЛОЖЕНИЙ СПЕЦИАЛЬНОЙ И ОБЩЕЙ ТЕОРИЙ ОТНОСИТЕЛЬНОСТИ В ФИЗИКЕ

Философское значение теории относительности

Современная западная цивилизация родилась в горниле череды религиозных и политических революций 16–17 веков. Но была ещё одна революция, которая определила духовное наполнение этой цивилизации – научная революция. Научная революция 17 века привела к рождению парадигмы классической науки. В основание этой парадигмы легла механика И.Ньютона.

Вторая научная революция зародилась в начале 20 века. И вновь у руля её – физика. В основание начала лежит парадигма релятивистских механики и электродинамики – специальная и общая теории относительности, а также квантовые механика и квантовая теория поля, - новая физика, которая ассоциируется с именами А. Пуанкаре, А. Эйнштейна, М. Планка и др. И в первом и во втором случае мы обнаруживаем две картины мира во многом отличных друг от друга. Картина мира – мировоззренческое, философское понятие. И. Ньютон прекрасно это понимал, ведь главный свой труд он наименовал «Математические начала натуральной философии» (1687). А. Эйнштейн также прекрасно понимал философское значение неньютоновской картины мира, о чём свидетельствует ряд его работ.

В обеих картинах мира ключевое место занимает связь понятий пространство – время – материя – движение

[1]. Причём, именно «время» в этой связке было самым загадочным звеном. С древних времён неуловимость сущности времени озадачивала мыслителей. Аристотель сформулировал «парадокс времени» - «прошлого уже нет, будущего ещё нет, значит, время – это только настоящее». На такое утверждение Аристотеля св. Августин остроумно ответил, что, «если настоящее в неуловимое мгновение становится прошлым, значит времени вообще нет». Конечно, св. Августин не сомневался в существовании времени, но у него не было ответа на вопрос – что есть время? «Что же такое время? Если никто меня об этом не спрашивает я знаю, что такое время; если бы я захотел объяснить спрашивающему – нет, не знаю». С трактовкой пространства у древних такой проблемы не было. Геометрическая его трёхмерность Евклидом и получила общественное признание.

Сэр Исаак Ньютон дал миру, как он считал, окончательное решение пространственно – временного строения мира. Согласно Ньютону, пространство и время предстали однородными, универсальными и абсолютными формами. Пространство и время – это объективные данности, существующие сами по себе, как вместилище всего существующего и длительность всего происходящего. Картина мира, предложенная И. Ньютоном, опирается на четыре принципа – «механицизм», «лапласовский детерминизм», «материализм», т.е. сведение материи к «веществу», как простому данному, «абсолютизм» (уверенность в абсолютной истинности ньютонианской механистической картины мира) [2]/

Вторая научная революция перечеркнула уверенность «ньютонианцев» в разгаданности строения мироздания.

Революционный переворот приводит к появлению неклассической физики, а затем, некоторое время спустя, и к парадигме неклассической (постклассической) науки. Такой переход мировым научно – философским сообществом

связывается с рождением специальной теории относительности (СТО) в 1905г. Но эта революция была подготовлена рядом открытий конца 19 века - делимости атома, определения природы электрона, введения постоянной Планка (квант действия) и, вместе с тем, закладыванием основ квантовой теории, открытием радиоактивности и т.п.

Что касается приоритета в создании СТО, то здесь, как правило, указывают Анри Пуанкаре и Альберта Эйнштейна. Как принято, рассмотрим хронологию основополагающих публикаций на тему СТО. Эту тему подробно исследовал академик А.А. Логунов [3]. В предисловии автор пишет: «Анри Пуанкаре уже в первой работе от 5 июня 1905 г. (Poincare H. Sur la dynamiq de l'electron // Comptes rendus hebdomadaires des séances de l'Akademie des scienes. Paris, 1905. V. 140. - P. 1504-1508), исходя из уравнений Максвелла – Лоренца, сформулировал принцип относительности для электромагнитных явлений как строгую математическую истину, а также открыл группу пространственно-временных преобразований и назвал её группой Лоренца, распространил преобразования Лоренца на все силы природы, независимо от их происхождения, в том числе на гравитационные.

В работе от 23 июля 1905 г. Он сформулировал всё основное, что является содержанием теории относительности, открыл законы релятивистской механики».

Первая и основная работа А. Эйнштейна по теории относительности «К электродинамике движущихся тел» поступила в редакцию 30 июня 1905 г. и была опубликована в журнале Zur Elektrodynamik der bewegter Körper. Ann. Phys, 1905, 17, 891-921.

Далее А.А. Логунов отмечает, что рассматриваемые статьи Пуанкаре – это изумительное по глубине и точности выражения великое творение величайшего естествоиспытателя. В них имеется почти всё основное, что составляет содержание теории относительности. В этом может убе-

диться каждый, кто имеет возможность внимательно изучить эти работы. В.Паули уже в зрелом возрасте оценил значение работ А.Пуанкаре и писал: «В совпадении результатов, полученных независимо друг от друга Эйнштейном и Пуанкаре, я усматриваю глубокий смысл гармонии математического метода и анализа, опирающегося на всю совокупность данных физического опыта».

Вместе с тем А. Пуанкаре дал науке и её философскому осмыслению важнейшее методологическое нововведение – принцип конвенционализма [4]. В связи с привлечением к описанию физических процессов неевклидовых геометрий он предложил характеризовать положения различных математических теорий как соглашения учёных принимать эти положения вне соображений их истинности или ложности. Этот принцип стал одним из основных в современной философии науки – все научные теории опираются на произвольные соглашения (конвенции) научного сообщества из соображений удобства, простоты, полезности этих теорий. Единственным ограничением выступает требование непротиворечивости трактовок. «Конвенционалисты», по мнению ведущего представителя логического позитивизма и философии науки Р. Карнапа [1], считали – «принять мир вещей значит лишь принять определённую форму языка». Опираясь на Витгенштейна и Рассела, Карнап считал предметом философии науки анализ структуры естественно–научного знания с помощью понятий математической логики. В последние годы жизни Карнап решительно высказывал положение о существовании «ненаблюдаемых материальных объектов» как основы для построения логических систем, близкое к естественно–научной материалистической тенденции. Радикальный конвенционализм К.Айдукевича определил, что отображение объектов в корне зависит от произвольного выбора понятийного аппарата (терминологии). Ещё ранее «неклассические» физики разошлись в понимании принципа дополнительности и принципа неопределённости – открыватели этих принци-

пов Н. Бор и В. Гейзенберг полагали, что они соответствуют объективным реалиям микромира, их оппоненты А. Эйнштейн и Э. Шрёдингер считали, что за этим стоит реальность сознания наблюдателя.

В свете сказанного, считаем возможным очень кратко остановиться на различных высказываниях по некоторым терминологическим понятиям и определениям философов различных поколений; остановиться на тех понятиях, которые прошли всю вертикаль в свете тематического анализа науки, в частности физики, и не утратили своего значения.

Диалектика – философская концептуализация развития, понятого как в онтологическом, так и в логико–понятийном его опрделениях - как в качестве теории, так и в качестве метода. По поводу претензии диалектики на универсальную всеохватность К. Поппер писал: «Любое развитие можно подогнать под диалектическую схему, и диалектика может не опасаться опровержения будущим опытом».

Материя – в материалистической традиции обозначает субстанцию, обладающую статусом первичности по отношению к сознанию. Последнее научно определил Менделеев: «Вещество, или материя, есть то, что наполняя пространство, имеет вес, т.е. представляет массы, то из чего состоят тела природы и с чем совершаются движения и явления природы». В современной философии очень немногие философы пользуются понятием «материя». Чаще естествоиспытатели говорят о материи, как о субстрактной первооснове вещей = вещество. Юлия Кристева, проводя диаматовский анализ практик сознания, говорит, что материя – то, что «не есть смысл», то «что есть без него, вне его и вопреки ему». Данная радикальная гетерогенность (материя / смысл) определяется также в качестве «поля противоречия» [5].

Время – форма возникновения, становления, течения, разрушения в мире, а также его самого вместе со всем

тем, что к нему относится. Объективное время – измеряется отрезками пути небесных тел. Субъективное время – осознание времени, зависящее от содержания переживаний. Время – способ существования человека – темпоральность безусловного условия бытия человека. В современной физике исходят из того, что объективного времени нет, оно – лишь координата в четырехмерном континууме. Вневременность мира означает,что событиями ра

зличных движущихся систем не может быть никакой одновременности: в материальном мире нет никакого времени. Мировая линия – это в континууме Римана кривая, изображающая движение точки в 4-мерной системе координат и т.п.[1]. Обобщая высказывания философов о пространстве и времени кратко их можно определить следующим образом: **пространство** – *совокупность отношений, выражающих координацию сосуществующих объектов - их расположение друг относительно друга и относительную величину (расстояния и ориентацию);* **время –** *совокупность отношений, выражающих координацию сменяющих друга состояний (явлений), - их последовательность и длительность* [1].

А.А. Зиновьев [6] пишет о том, что «физические теории пространства и времени не дают определений пространственно – временной терминологии в том смысле, что не вводят эту терминологию впервые в употребление через другие термины с известным значением. Они предполагают эту терминологию данной и формулируют методы установления пространственно – временных отношений предметов для различного рода случаев, в частности для различно движущихся систем, для удалённых событий, для случаев, когда имеет значение скорость распространения сигналов о событиях и т.д. Для этих целей физические теории осуществляют *экспликацию* пространственно – временной терминологии».

Экспликация пространственно – временных характеристик в физике проведена, опираясь на работу [6] и

статью [7]. Здесь лишь напомним, что измерение физической величины сводится к сравнению её с другой однородной величиной, точно определённой и условно принятой за единицу (*эталон*) для измерения всех таких же величин. Поэтому для пространственно – временных (П-В) измерений необходимо перейти от философского определения П и В к конкретно – физическому («измеренческому»), т.е. выраженных при помощи определённых физических объектов (линейки, часы, спектрографы и т.п.) и процессов (*экспликация понятий*). Отметим также, что основные П-В эталоны длины и времени определены физически одинаково – по спектрам излучения атомов (криптон-86 и цезий-133) с учётом ограничения предельной скорости передачи информации и взаимодействия объектов. Такой фундаментальной физической постоянной скорости является скорость движения *безмассовых частиц*, в частности – фотонов. Действительно, передача информации (воздействие одних объектов на другие) при помощи определённого носителя принципиально не может быть бесконечно большой, т.к. такой *носитель*, в этом случае, мог бы одновременно находиться в любом месте Вселенной, т.е. «быть везде».

Мир познаваем вследствие одинаковости законов физики во всех «уголках» пространства Вселенной и во все времена. Это обстоятельство позволяет познавать события, происшедшие миллиарды лет назад, на громадном удалении от Земли, даже когда ни Земли, ни Солнечной системы ещё не было, и прогнозировать развитие физических процессов в будущем. Спектральный состав и интенсивности линий электромагнитного излучения атомов, молекул, конденсированных объектов и т.п., позволяют понять многие процессы, происходящие в макро- и микромире.

Условие познаваемости происходящих в Мире событий приводят к равноправию различных систем отсчёта (СО), как инерциальных (ИСО), так и неинерциальных (НИСО – точнее локально ИСО).

Несколько слов о соотношении *абсолютного и относительного* в событиях, происходящих и рассматриваемых в различных СО. С этой точки зрения любое физическое событие (излучение атомов, молекул, радиоактивных ядер, взрыв сверхновой звезды и т.п.) произошедшее в одной из СО, произошло **абсолютно**. Другое дело, информация об этом событии в различных СО будет «выглядеть» по-разному, будет **относительной**. Эта относительность описания одного и того же события в разных СО обусловлена характером (направлением и величиной) относительной скорости между источником информации (ИИ) и приёмником информации (ПИ) вследствие эффекта Доплера. Открытие Пуанкаре группы Лоренца и фундаментального инварианта $\Sigma^4_{i=1} x^2_i = inv$ позволили ему построить ряд четырёхмерных величин, которые при лоренцевых преобразованиях изменяются как время и пространственные координаты [3]. Поэтому различные наблюдатели, изучающие рассматриваемое событие в «своих СО», используя преобразования Лоренца, получат одинаковые сведения (абсолютные) о произошедшем событии.

События, происходящие в П-В, могут быть связаны отношением «*причина - следствие*», а могут в принципе не быть связаны таким отношением. Действительно, если скорость v передачи «*воздействия*» события А (причина) на событие В (следствие) конечна, расстояние АВ = Δx, время распространения сигнала от А до В составляет Δt , то события А и В могут быть связаны отношением причина – следствие лишь при условии $v\Delta t \geq \Delta x$. Поскольку максимальная скорость передачи информации – скорость света вакууме c, то событие А может причинить событие В лишь при условии, что интервал между ними $\Delta S = c\Delta t - \Delta x \geq 0$ Такой интервал между событиями А и В называют *времениподобным,* в случае, если $\Delta S < 0$ – *пространственноподобным.* Причинно – следственная связь возможна лишь для событий, связанных времениподобным интервалом.

Эффект Доплера и общая теория относительности.

Рассмотрим изменение спектра излучения (ω_0, T_0, λ_0) некоторого источника излучения ИИ принимаемого сигнала приёмником излучения – ПИ (ω, T, λ) вследствие эффекта Доплера. Пусть ИИ движется относительно ПИ вдоль линии, их соединяющей, с ускорением a. Такое движение может происходить при удалении, например, космического корабля (ИИ), или при свободном падении ИИ в гравитационном поле некоторого аттрактора – А (звезды, планеты, если угодно, - «*чёрной дыры»* и т.п.*)*. Спектры, принятые ПИ, смещены относительно спектра ИИ. Это смещение будет зависеть от их относительной скорости в каждый момент времени, которая, в свою очередь, зависит от ускорения ($v = v_0 + a\tau$). Подчеркнём, что численные значения отношений частот, или длин волн, спектральных линий ИИ и ПИ остаётся постоянным, что даёт возможность идентифицировать природу источника излучения и характер его движения относительно приёмника этого излучения.

Период принимаемой волны каждой линии спектра, определяемый методом k – коэффициентов [7], по периоду излучения имеет вид:

$$\sqrt{\frac{1+\frac{a\tau}{c+v_0}}{1-\frac{a\tau}{c-v_0}}}$$

$$\mathbf{T = T_0 \sqrt{\{[c + (v_0 + a\tau)]/[c - (v_0 + a\tau)]\}} = T_0 \sqrt{[(c + v_0)/(c - v_0)]},}$$

Здесь $\tau = NT_0$, N = 1, 2, 3, …, r/λ_0, $\lambda_0 = cT_0$, $\tau = r/c$, где r – расстояние между ИИ и А.

Выполним подстановку $ar/c(c+v_0) = \alpha$, $ar/c(c-v_0) = \delta$ и рассмотрим второй корень, учитывая, что

$$\alpha, \delta << 1: \sqrt{[(1+\alpha)/(1-\delta)]} \cong \left(1+\frac{\alpha}{2}\right)\left(1+\frac{\delta}{2}\right) \cong 1 + \frac{ar}{c^2}$$

и тогда

$$T = T_0 \sqrt{\frac{c + v_0}{c - v_0}} \cdot \frac{1}{\sqrt{1 - \frac{2ar}{c^2}}} .$$

Аналогично,

$$\omega = \omega_0 \sqrt{\frac{c - v_0}{c + v_0}} \cdot \sqrt{1 - \frac{2ar}{c^2}} .$$

Если ИИ движется в поле А только под действием его силы тяготения, то ускорение $a = g = GM / r^2$ – напряженности гравитационного поля аттрактора.

В ОТО величину $2GM / c^2 = r_g$ называют гравитационным радиусом аттрактора.

Таким образом, можно записать:

$$T = T_0 \sqrt{[(1+\beta) / (1-\beta)]} \frac{1}{\sqrt{1 - \frac{r_g}{r}}}$$

и для частоты

$$\omega = \omega_0 \sqrt{\frac{1 + \beta}{1 - \beta}} \cdot \sqrt{1 - \frac{r_g}{r}} .$$

Рассмотренные изменения частоты, периода и длины волны полностью укладываются в рамки эффекта Доплера. Однако, в статье Я.Б. Зельдовича и И.Д. Новикова «Общая теория относительности и астрофизика» в «Эйнштейновском сборнике» за 1966 г. по поводу этой формулы сказано следующее: «второй множитель после ω_0 описывает замедление времени в гравитационном поле, первый – эффект Доплера». Воздействием гравитационного поля на γ – фотоны Паунд и Ребка объяснили гравитационное красное смещение частоты фотонов в эффекте Мёссбауэра при прохождении ими гравитационной разности потенциалов $\Delta\varphi$:

$$\Delta\omega/\omega = \Delta\varphi/c^2 = gL/c^2.$$

Здесь L – расстояние по вертикали между ИИ и ПИ, причем здесь принято, что масса фотона $m_\phi = h\omega / 2\pi c^2$. Учи-

тывая, что $m_ф = 0$, остаётся предположить воздействие гравитационного поля на энергетическую структуру нуклонов в ядрах атомов ИИ и ПИ.

Для того, чтобы убедиться в возможности такого воздействия, можно предложить проведение следующего эксперимента. На горизонтальном столе, в одном уровне ($\Delta\varphi = 0$) находятся ИИ и ПИ. Пусть они будут такими же, как в опыте Паунда и Ребки. ИИ прикрепить к дну цилиндрической керамической трубки, расположенной горизонтально. Во внутрь трубки ввести один спай дифференциальной термопары. Другой спай присоединить к керамическому держателю пластинки - поглотителю γ – фотонов, за которым находится детектор γ – излучения. Поскольку ИИ и ПИ находятся в одном уровне и и их относительное движение отсутствует, то исключаются влияния и гравитационного поля и эффекта Доплера.

Для исключения возможного влияния на энергетические состояния ядер атомов ИИ магнитного поля нагревающая обмотка на поверхности трубки должна быть бифилярной. Теперь осталось убедиться, в том, что при разности температур ИИ и ПИ $\Delta T = 0$ наблюдается резонансное поглощение поглотителем γ – фотонов, испущенных ИИ. При нагреве ИИ до температуры выше температуры ПИ ($3k\Delta T / 2$) на такую величину, что частота γ – лучей ИИ в достаточной степени изменится по сравнению с резонансной частотой поглощения ПИ. Ясно, что изменение теплового воздействия на энергетическую структуру ядер ИИ должно быть такого же порядка, как и при изменении энергии в опытах Паунда и Ребки, или при изменении частоты вследствие эффекта Доплера. Конечно, должны быть соблюдены правила техники безопасности при работе с радиоактивными элементами.

Релятивистский закон всемирного тяготения

Применим понятие «сцеплённые состояния» [8] к выводу формулы релятивистского закона всемирного тяготения (РЗВТ). В работе [7] было отмечено, что масса тела является скалярной величиной, не зависящей от скорости его движения относительно любой системы отсчёта. Вследствие этого возможны два пути вывода РЗВТ:

1. Записать выражение для релятивистского интервала (x_i, i = 1, 2, 3, 4) для движущегося тела, взять последовательно две производные по собственному времени и получить выражение для 4-вектора ускорения. Произведение массы тела на это ускорение даст выражение для 4-вектора силы.

2. Записать выражение для 4-вектора импульса в инерциальной системе отсчёта, взять от него производную по собственному времени – вновь получим 4-ветор силы. Такие операции были проделаны в [8]. Далее выполним «сцепление», аналогичное [9], выражения для релятивистской силы с ЗВТ и получим РЗВТ.

Рассмотрим два случая движения тела в центральном гравитационном поле некоторого аттрактора (А). Пусть в первом случае импульс тела изменяется только по направлению, т.е. сила направлена перпендикулярно к импульсу ($\overrightarrow{\mathbf{F}} \perp \vec{\mathbf{p}}$) – импульс тела по величине не меняется. Тогда

$$\overrightarrow{F_{\perp}} = \frac{m\vec{a}}{\sqrt{1-\frac{v^2}{c^2}}}.$$

Во втором случае сила параллельна импульсу ($\vec{F} \; II \; \vec{p}$). Тогда

$$\overrightarrow{F_{II}} = \frac{m\vec{a}}{\left(1-\frac{v^2}{c^2}\right)^{\frac{3}{2}}}.$$

Учитывая, что ускорение тела в гравитационном по-

ле (иначе, только под действием сил тяготения) совпадает с численным значением напряженности гравитационного поля в точке мгновенного нахождения тела ($\vec{a} = \vec{g}$), можно утверждать, что

$$a = G\frac{M}{r^{2,}}, \quad v^2 = G\frac{M}{r}$$

в первом случае, и

$$F_1(= G\ Mm/(r^{\uparrow}2\ \sqrt{(1 - GM/(c^{\intercal}2\ r))})$$

во втором случае (при радиальном падении).

$$\frac{v^2}{2} = G\frac{M}{r}\ .$$

Сила, действующая на тело будет равна (второй случай):

$$F_{II} = G\frac{Mm}{r^2\left(1 - \frac{2GM}{c^2 r}\right)^{\frac{3}{2}}}\ .$$

В физике принято обозначение $\frac{GM}{c^2} = r_g$ – гравитационный радиус. Так что обе формулы могут быть записаны в следующем виде:

$$F_{\perp} = G\frac{Mm}{r^2\sqrt{1 - \frac{r_g}{r}}} \quad \text{и} \quad F_{II} = G\frac{Mm}{r^2\left(1 - \frac{2r_g}{r}\right)^{\frac{3}{2}}}.$$

Здесь G – гравитационная постоянная, M – масса аттрактора, m – масса тела, r – расстояние между телом и аттрактором, c – скорость света, v – скорость тела.

Из приведенных формул видно, что при сближении тела с аттрактором на расстояние приближающееся к гравитационному радиусу сила их взаимодействия (притяжения) неограниченно растёт вплоть до бесконечно больших значений. Такие объекты в астрофизике называют «чёрны-

ми дырами».

В заключении следует отметить, что понимание специальной и общей теории относительности требует глубинного осознания таких категорий как «пространство», «время» и формирование у исследователя некоторого «физического» мировоззрения. Ключевым пунктом этого понимания является одинаковость законов физики в любых системах отсчета, из чего следует принципиальная познаваемость мира, основанная на регистрации и анализе физических сигналов, исходящих от исследуемого объекта. При таком подходе теория относительности фактически играет роль метода коррекции физических сигналов на величину скорости движения систем отсчета и наличие гравитации, позволяя избежать искажения информации в реальном физическом эксперименте. Какие-либо «парадоксы» теории относительности не несут физического смысла и являются, по нашему мнению, следствием непонимания ее основ.

Список литературы

1. Современный философский словарь. Лондон; Франкфурт-на-Майне; Париж; Люксембург; Москва; Минск.: ПАНПРИНТ, 1998. – 1064 с.
2. Шенгелая І.Ш. Історія філософії. - Севастополь: Вид.-во СевНТУ, 2012. – 372 с.
3. Логунов А.А. К работам Анри Пуанкаре «О динамике электрона». 2-е изд. перераб. и доп. – М.: Изд-во МГУ, 1988. – 103 с.
4. Пуанкаре Анри. О науке: Пер. с франц. – М. Наука. Гл. ред. Физ.-мат. литературы. 1983. – 560 с.
5. История философии: Энциклопедия. - Мн.: Книжный Дом, 2002. – 1376 с.
6. Зиновьев А.А. Логика науки.- М.: «Мысль». 1971. – 2798 с.
7. Леонтович А.Л., Евстигнеев М.П. Нетрадиционное изложение специальной теории относительности. Вестник

СевНТУ. Вып. 99: Физика и математика. Сб. науч. Работ. – Севастополь: Изд. СевНТУ, 2009. – 156 с.
8. Леонтович А.Л. К вопросу о «сцеплённых состояниях». Материалы 6-й Международной научно-технической конференции «Актуальные вопросы теоретической и прикладной биофизики, физики и химии», т.1, с.150-154. БФФХ – 2010 – изд. СевНТУ.
9. Леонтович О.Л., Євстигнєєв М.П. До формалізму спеціальної і загальної теорій відносності. Материалы 7-й Международной научно технической конференции «Актуальные вопросы БФФХ», с. 379-382. БФФХ-2011- Изд. СевНТУ.

УДК 517.977.5; 681.5.03
А.Б. Филимонов[1], д-р техн. наук, проф.
Н.Б. Филимонов[2], д-р техн. наук, проф.
[1]*МАТИ - РГТУ им. К.Э. Циолковского, г. Москва, Россия*
[2]*МГУ им. М.В. Ломоносова, ИПУ им. В.А. Трапезникова РАН, г. Москва, Россия*
nbfilimonov@mail.ru

СИТУАЦИОННО-ЦЕЛЕВОЙ ПОДХОД В ЗАДАЧАХ АВТОМАТИЗАЦИИ ЭРГАТИЧЕСКИХ СИСТЕМ УПРАВЛЕНИЯ ПОДВИЖНЫМИ ОБЪЕКТАМИ

В современном техногенном обществе важную роль играют эргатические системы управления (ЭСУ) подвижными объектами (ПО), т.е. целеустремленные системы «человек–машина–среда» [1]. В данных системах человек-опе-ратор (ЧО), взаимодействуя с техническими средствами, выполняет функции целеполагающего характера и осуществляет управление движением объекта в ручном и полуавтоматическом или супервизорном режимах. К данному классу автоматизированных систем относятся управляемые человеком автомобильные, железнодорожные, водные и воздушные транспортные средства, а также системы дистанционного управления манипуляционными роботами, радиолокационными станциями, системы телеуправления и т.п.

Общую тенденцию развития ЭСУ ПО характеризуют следующие факторы: возрастание степени сложности задач управления, повышение степени потенциальной опасности внешней среды, рост объема обрабатываемой информации, сокращение времени на принятие решений, возрастание цены ошибок, правильности и качества принимаемых оперативных решений и др. Работа оператора в этих условиях чревата сенсорными, эмоциональными и интеллектуальными перегрузками [2-4]. В связи с этим необ-

ходима новая идеология автоматизации ЭСУ ПО, тированная на оптимальное распределении функций между оператором и автоматикой. В докладе обсуждается развиваемая авторами методология автоматизации эргатических систем данного класса, основанная на иерархической декомпозиции функциональных задач, центральное место в которой занимает концепция ситуационно-целевого управления.

Структура, режимы и степени автоматизации ЭСУ ПО

Функциональная структура ЭСУ ПО схематично представлена на рисунке 1. Здесь средства автоматизации (СА) реализуют процессы обработки информации, контроля и управления. ЧО воспринимает и перерабатывает инструментальную, получаемую от средств отображения информации (СОИ), и неинструментальную, получаемую от рецепторов органов чувств, информацию об окружающей среде. Человеко-машинный интерфейс (ЧМИ), образованный СОИ и органами управления (ОУ), обеспечивает взаимодействие оператора с автоматикой.

В процессе деятельности ЧО выполняет следующие функции: сенсорные - прием информации, интеллектуальные - восприятие, оценка и переработка информации, волевые - принятие решения, эффекторные - реализация решения. Психика, психические процессы являются регуляторами деятельности, посредством которых оператор направляет, контролирует и корректирует свои действия в соответствии с поставленной целью.

Важнейшим свойством ЭСУ ПО является *многорежимность*. В зависимости от степени автоматизации возможны различные режимы управления движением: *ручное* (осуществляемое непосредственно оператором), *полуавтоматическое* (осуществляемое автоматикой с участием оператора), *автоматическое* (осуществляемое автоматикой без участия оператора). В условиях компьютеризации ЭСУ ПО необходимо учитывать два способа участия ЧО в про-

цессах автоматизированного управления - *диалоговый* (интерактивный) и *супервизорный* режимы: в первом случае осуществляется непосредственное взаимодействие оператора и автоматики, а во втором - оператор осуществляет наблюдение за работой автоматики и ограничивается лишь оперативным вмешательством с целью ее коррекции.

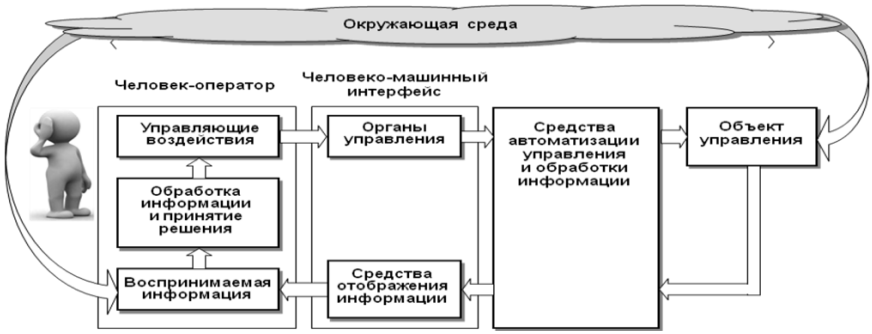


Рисунок 1

Иерархия функциональных задач ЭСУ ПО

Эффективное решение задач автоматизации ЭСУ ПО возможно лишь в рамках комплексного интегрированного подхода, связанного с кардинальным изменением функциональной структуры системы - расширением состава автоматизируемых функций.

На рисунке 2 представлена предложенная авторами трехуровневая структура автоматизации ЭСУ ПО [5, 6]. Здесь на верхнем уровне планируются сценарии выполнения задания, на среднем уровне осуществляются функции ситуационно-целевого управления, на нижнем уровне вырабатываются управляющие воздействия на объект.

Функционирование ЭСУ ПО подчиняется конкретному ***заданию***, которое определяет оперативную цель действий системы.

При выполнении системой задания ЧО руково-

дствуется некоторым оперативным планом действий, торый следует относить к ***сценарному планированию*** [7] - разработке альтернативных сценариев будущего и создании на их основе гибких оперативных планов действий. Под сценарием понимается упорядоченная последовательность сцен, которые определяются стереотипной ситуацией, сложившейся вследствие наступления некоторого набора событий, и действиями в данной ситуации. Сценарий разворачивается в процессе функционирования системы, причем все альтернативные сценарии формируются на основе исследования окружающей среды (наличие предопределенных элементов, ключевых неопределенностей и их комбинации). Сценарное планирование рассматривает все сценарии как одинаково возможные в будущем и придает гибкость стратегиям управления, помогая сократить время реагирования на изменение ситуаций.

Отработка плана осуществляется контуром ***ситуационно-целевого управления*** (СЦУ), который реализуется алгоритмически (в результате человеко-машинного взаимодействия) и сводится к последовательному выполнению следующих трех функций: мониторинг ситуаций; генерация локальной цели уп-равления; формирование стратегии управления. Здесь базисным является понятие *ситуации* - конкретного набора событий, обстоятельств и условий, которые влияют на систему в данное конкретное время. Все три задачи СЦУ вытекают из логики отработки сценария. Решением данных задач являются соответственно диагноз текущей ситуации, новая локальная цель управления (в случае изменения ситуации) и стратегия управления, направленная на ее достижение.

Уровень ***управления движением*** ПО является нижним уровнем функциональной иерархии ЭСУ и реализует стратегии управления, сформированные на верхних уровнях, в автоматическом или автоматизированном (диалоговом или супервизорном) режиме.

Рисунок 2

Средний и верхний уровни направлены на автоматизацию эргатических функций и их проектирование выходит за рамки традиционных схем автоматизации. Ключевую роль в автоматизации ЭСУ ПО является средний иерархический уровень - уровень СЦУ. Остановимся подробнее на его специфике.

Особенности ситуационно-целевого управления

Ситуационная методология - чрезвычайно многогранное и многоплановое направление, получившее в настоящее время довольно широкое распространение в самых различных научных сферах и областях деятельности: технике, социологии, психологии, педагогике и др. Основное

значение ситуационного подхода в науке состоит в попытпытке теоретически сформулировать, эмпирически проверить и затем практически рекомендовать различные решения применительно к каждой из типовых ситуаций. Диагностика ситуации позволяет определить альтернативные алгоритмы по ее разрешению, а учет реальных субъективных и объективных факторов, влияющих на ситуацию, дает возможность выбрать наиболее оптимальный путь по ее разрешению.

Ситуационный подход в отличие от других, ранее претендовавших на роль универсальных и «единственно верных», показывает, что лучшего способа управления нет в принципе: эффективность любого из них относительна и определяется текущей ситуацией. Ситуационный подход можно охарактеризовать как концепцию «управленческой относительности», резко контрастирующую с абсолютизмом многих иных подходов и их претензиями на универсальность. Неслучайно, переход от «универсализма» к «ситуационизму» сравнивают с «переходом от плоскости к трехмерному пространству, от немого черно-белого кино к цветному со стереоскопическим звуком».

Ситуационная методология имеет давнюю историю и берет свое начало в теоретическом менеджменте. Еще в начале 1920-х гг. известный социолог Фоллетт (M.P. Follett) сформулировала «закон ситуации», согласно которому эффективность управленца зависит от обладания им необходимыми знаниями, причем «различные ситуации требуют различных типов знаний». Укажем также на ситуационную модель лидерства, разработанную известным организационным психологом Фидлером (F.E. Fiedler) в середине 1950-х гг., согласно которой эффективность руководства зависит от конкретной ситуации: для разных ситуаций требуются различные типы лидеров, а тип идеального лидера отсутствует.

В кибернетических системах ситуационный подход сформулирован в работах 1963-1969 гг. одного из отцов

искусственного интеллекта Маккарти (J. McCarty) [8, 9]. Он разработал способ описания модели мира и ее изменений в рамках формализма математической логики и ввел понятие *ситуационного исчисления* (Situational Calculus) - формального аппарата для построения модели мира и выполнения рассуждений о воздействиях на мир. Позднее были разработаны альтернативные подходы: ситуационное исчисление Рейтера (R. Reiter), исчисление событий Миллера (R. Miller) и Шенхема (M. Shanahan) и др.

Существенный вклад в развитие ситуационного подхода внесли работы отечественные ученых в начале 1970-х гг. Здесь прежде всего необходимо выделить модельную теорию мышления В.Н. Пушкина [10] и метод ситуационного моделирования Ю.И. Клыкова и Д.А. Поспелова [11]. На их основе были разработаны методы ситуационного управления большими системами [12, 13].

Центральным понятием ситуационного управления является управленческая ситуация. Само понятие *ситуации* используется довольно часто в самых разнообразных аспектах и порой неотделимо от таких понятий как состояние, событие, процесс, положение и т.д. Так, например, Ю.И. Клыков и Д.А. Поспелов в своих работах явно отождествляют ситуацию с состоянием.

Успех ситуационного подхода в значительной мере зависит от точного определения наиболее значимых факторов, влияющих на процесс управления. Классическое ситуационное управление подчиняется схеме «ситуация–дейст-вие», причем алгоритм управления явно не задается, а реализуется на ситуационной сети. Под *ситуационной сетью* понимается ориентированный взвешенный граф переходов по эталонным ситуациям, определяемым исходной и целевой ситуациями. Дуги графа отражают смену ситуаций и вызывающие ее информационно-управляющие процессы. Заметим, что понятие дискретной сети проблемных ситуаций было введено В.Н. Пушкиным и Д.А. Поспеловым.

В отличие от классического ситуационного подхода, воплощаемого двухступенчатой схемой «ситуация–действие», в рассматриваемой функциональной структуре ЭСУ ПО подразумевается четырехступенчатая схема «ситуация – цель – стратегия управления – действие», иллюстрируемая на рисунке 3.

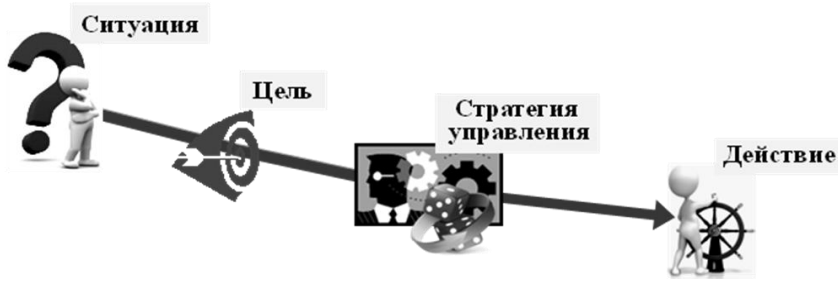


Рисунок 3

Развиваемая авторами концепция ситуационного управления имеет существенные отличия от существующих подходов и характеризуется следующими базовыми положениями [14]:

1. Ситуация не является совокупностью переменных состояния объекта и внешней среды, а идентифицирует некоторую область локализации динамических переменных. К примеру, при малых возмущениях текущего динамического режима состояние системы изменяется, а ситуация остается низменной.

2. Каждая ситуация порождает локальную (рабочую) цель управления, для достижения которой формируется стратегия управления. Тем самым ситуационное управление по своей структуре является дискретным.

3. Предлагается использовать классификационные схемы для ситуаций, в частности, выделять критические (предаварийные и аварийные) ситуации. На этой основе решается задача ситуационного мониторинга.

4. Предлагается ситуационный подход сочетать с многоцелевым принципом управления сложными системами: главная цель разбивается на подцели, которые формализуются как локальные цели управления применительно к отдельным ситуациям. В итоге формируемые стратегии являются прогностическими, учитывающими возможные сценарии ожидаемого будущего развития ситуации.

5. Излагаемая трактовка ситуации позволяет характеризовать и анализировать динамические процессы в объекте с учетом различных факторов неопределенности в динамике объекта и свойствах окружающей среды.

6. Ситуационное пространство динамических объектов является конечным и представляется ситуационным графом. Тем самым открывается возможность построения аппарата ситуационного исчисления и разработки логико-динамического формализма ситуационно-целевого управления.

Ситуационное моделирование процессов управления

Ситуационно-целевое управление базируется на *ситуационной модели*, дающей формальное описание мира с помощью ситуаций, в которых предстоит действовать ЭСУ. В модели учитываются те факторы, которые действительно влияют на развитие ситуаций. Сюда относятся релевантные характеристики объекта и внешнего мира: пространственное положение, режим и условия движения объекта, его техническое и динамическое состояние.

Необходимой предпосылкой эффективности метода ситуационно-целево-го управления является возможность различного уровня абстрагирования при построении ситуационных моделей реальных процессов. Такую возможность предоставляет описываемая ниже конструкция ситуационного пространства - в ней определены операции агрегирования и дробления ситуаций.

Ситуационное пространство определяется парой (Ω, S), где Ω - носитель ситуационного пространства, $S \subset \Omega$

- семейство его подмножеств, именуемых *ситуациями*. Будем интерпретировать Ω как пространство *элементарных* ситуаций $\omega \in \Omega$. Тогда конкретная ситуация s - некоторая сумма элементарных ситуаций: $s \subset \Omega$. Таким образом, к ситуациям применимы обычные теоретико-множественные операции. Операция объединения приводит к агрегированию ситуаций, а разбиения множеств - к дроблению ситуаций.

Отношение включения $s_1 \supset s_2$ порождает *дерево*, или *иерархию ситуаций*.

Древовидную структуру иерархии ситуаций можно представить т.н. моделью вложенных множеств (Nested Set Model) как показано на рисунке 4.

Операции агрегирования и иерархической структуризации ситуаций являются удобными инструментами, допускающими те же самые операции для аналогичной структуризации множества локальных целей управления.

Ситуационное моделирование на базе сетей Петри. Один из возможных вариантов развития ситуационного подхода - формализация динамических ситуационных моделей посредством аппарата сетей Петри.

Сеть Петри определяется кортежем

$$N = \langle P, T, F, \mu_0 \rangle,$$

где

- $P = \{p_1, p_2, \ldots, p_m\}$ - множество позиций;
- $T = \{t_1, t_2, \ldots, t_n\}$ - множество переходов;
- $F = (P \times T) \cup (T \times P) \rightarrow Z_+ = \{0, 1, 2, \ldots\}$ - функция инцидентности;
- μ_0: $P \rightarrow Z_+$ - начальная маркировка сети.

Последовательность срабатываний переходов $\tau = t_1, t_2, \ldots, t_l$ задает *динамику* сети Петри: μ_0, μ_l.

Динамическая модель системы получается на основе интерпретации позиций p как ситуаций, а переходов t как событий, инициирующих изменения ситуаций.

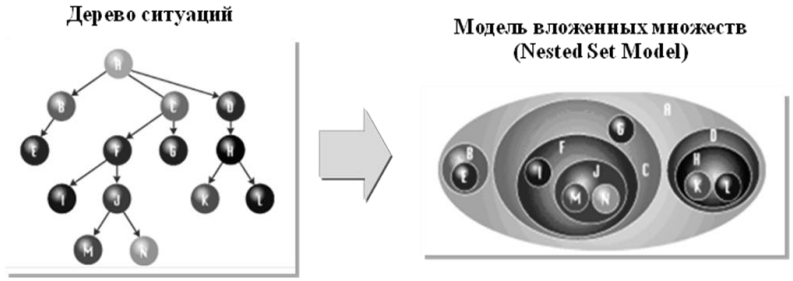


Рисунок 4

Ситуационное моделирование на базе сетей фреймов. В качестве одного из способов формализации ситуационных моделей предлагается аппарат фреймовых сетей.

Фреймовая модель, основанная на теории М.Минского, представляет собой систематизированную в виде единой теории технологическую модель памяти человека и его сознания. Фрейм - структура данных для представления некоторого концептуального объекта. Относящаяся к фрейму информация содержится в его компонентах - слотах.

Так фрейм ТЕКУЩАЯ СИТУАЦИЯ должен иметь следующие слоты:

- состояние объекта;
- состояние среды;
- глобальная цель;
- текущая локальная цель;
- новая локальная цель;
- стратегия управления.

В отличие от моделей других типов, во фреймовых моделях фиксируется жесткая структура, которая называется фреймом-прототипом, или образцом.

Значением слота могут быть данные различных типов. В частности, это могут быть наборы слотов более низкого уровня. В качестве значения слота может выступать

имя другого фрейма - так образуются сети фреймов. Все фреймы взаимосвязаны и образуют единую фреймовую структуру, в которой органически объединены декларативные и процедурные знания. Это дает возможность композиции и декомпозиции информационных структур.

Фреймовые модели является достаточно универсальными, поскольку позволяют отобразить все многообразие знаний о мире: через фреймы-структуры для обозначений объектов и понятий (цель); фреймы-роли (оператор); фреймы-сценарии (режим движения); фреймы-ситуации (авария, штатный режим).

Список литературы

1. Человеко-машинные системы управления / Машиностроение. Энциклопедия. Т. I-4. Автоматическое управление. Теория. Раздел 8 / Под общ. ред. Е.А.Федосова. - М.: Машиностроение, 2000. - С. 604-641.
2. Таран В.А. Эргатические системы управления. - М.: Машиностроение, 1976.
3. Шеридан Т.Б., Феррел У.Р. Системы человек-машина. Модели обработки информации, управления и принятия решений человеком-оператором. - М.: Машиностроение, 1980.
4. Павлов В.В. Синтез стратегий в человеко-машинных системах. - К.: Вища школа, 1989.
5. Теряев Е.Д., Петрин К.В., Филимонов А.Б., Филимонов Н.Б. Проблемы и перспективы автоматизации эргатических систем управления подвижными объектами / Мехатроника и робототехника: Итоговый сб. статей. - СПб.: Изд-во «Политехника-сервис», 2010. - С. 58-62.
6. Теряев Е.Д., Петрин К.В., Филимонов А.Б., Филимонов Н.Б. Современные проблемы автоматизации и интеллектуализации эргатических систем управления подвижными объектами / Интеллектуальные системы управления: Коллективная монография. Под ред. С.Н.Васильева - М.: Машиностроение, 2010. - С. 84-95.

7. Линдгрен М., Бандхольд Х. Сценарное планирование. Связь между будущим и стратегией. - М.: Изд-во «Олимп-Бизнес», 2009.
8. McCarthy J. Situations, Actions and Causal Laws // Stanford Artificial Intelligence. Technical report: Memo 2. Stanford University, 1963.
9. McCarthy J., Hayes P.J. Some Philosophical Problems from the Standpoint of Artificial Intelligence / Machine Intelligence 4. Edinburg: Edinburgh University Press, 1969. - P. 463-502.
10. Поспелов Д.А., Пушкин В.Н. Мышление и автоматы. - М..: Советское радио, 1972.
11. Клыков Ю.И. Семиотические основы ситуационного управления. - М.: МИФИ, 1974.
12. Клыков Ю.И. Ситуационное управление большими системами. - М.: Энергия, 1974.
13. Поспелов Д.А. Ситуационное управление. Теория и практика. - М.: Наука, 1986.
14. Чекинов Г.П., Чекинов С.Г. Ситуационное управление: состояние и перспективы // Приложение к журналу «Информационные технологии». - 2004. - № 2. - 32 с.

УДК 62-51
С.Ф. Сергеев, д-р психол. наук, проф.
Санкт-Петербургский государственный университет, ОАО «Корпорация «Аэрокосмическое оборудование», г. Санкт-Петербург, Россия
ssfpost@mail.ru

ПРОБЛЕМА ЭФФЕКТИВНОГО ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА-ОПЕРАТОРА С ИНТЕЛЛЕКТУАЛЬНЫМИ ТЕХНИЧЕСКИМИ СИСТЕМАМИ И СРЕДАМИ

Введение

Классические эргатические системы погружают оператора во взаимодействие с системой управления в рамках выполнения фиксированных алгоритмических действий с заранее ясными и наблюдаемыми следствиями в поведении управляемой системы [1]. Основная проблема согласования возникающих взаимоотношений лежит в области формирования у оператора соответствующей цели системы концептуальной модели и навыков управления с помощью органов управления. Решение данной задачи лежит в области психологии обучения и решается методами тренажерной подготовки на моделях реальной деятельности [2]. Многократное повторение профессиональной задачи приводит к выбору и усвоению наиболее эффективного для данного оператора в данной деятельности способа управления и формирует оптимальную форму профессионального поведения. По окончании обучения задача обеспечения эффективной связи «человек-машина» заключается в поддержании навыков оператора на требуемом уровне. Однако такой подход плохо работает в случае, если техническая система реализует алгоритмы искусственного интеллекта, демонстрируя оператору поведение, интерпретируемое им как интеллектуальное или включенное в соци-

альный контекст.

Информационные технологии предоставляют разработчикам эргатических систем широкий спектр средств, повышающих их интеллектуальность, формируя отношения человека-оператора с технической системой подобные отношениям, возникающим в условиях естественной социальной коммуникации. Вместе с тем взаимодействие человека с искусственными информационными средами, наделенными искусственным интеллектом отличается от его взаимодействия с естественными средами в силу структурно-функциональной дополнительности искусственных сред с когнитивными структурами человека [3]. Возникает проблема симбиоза между системами разной природы – биологическими, наделенными механизмами психического отражения и активного целеполагания, и техническими, реализующими технологии искусственного интеллекта.

Наиболее часто эффективность функционирования сложных систем связывают с понятием «системный гибридный интеллект» [4], которое получает новые смыслы при включении человека в сложноорганизованные среды. В определении автора данной концепции В.Ф. Венды гибридный интеллект рассматривается как «наиболее общее понятие, обозначающее совместное прогнозирование, в котором участвуют многие индивиды, располагающие частными моделями процессов взаимной адаптации» [4, с. 179]. Особое значение гибридных форм интеллектных объединений имеют техногенные среды эргатических систем объединяющие множество пользователей в рамках решения задач контроля и управления.

Техногенные среды ведут к новым формам интеграции человека с машиной, формируя гибридные и симбиотические формы естественного интеллекта, отличающиеся от классического интеллекта, возникшего в процессе жизнедеятельности человека в изолированных культурных средах и естественной среде жизнедеятельности. Это ведет к постановке задачи повышения его эффективности в про-

цессе взаимодействия с компонентами искусственного интеллекта эргатической системы. Цель настоящей статьи – привлечь внимание разработчиков эргатических систем к рассмотрению некоторых возможных вариантов симбиотического объединения интеллекта человека-оператора с интеллектом технических систем и сред.

Системные представления об интеллекте и интеллектных симбионтах

Поиски способов усиления возможностей естественного интеллекта, а в более широком плане – человеческого ума, возможностей создания их искусственных аналогов, приобретают особое значение в связи с интенсивным развитием техники и технологий и пониманием ограниченности индивидуального человеческого разума при решении задач управления [5]. Развитие информационных технологий питает наблюдаемый в последнее десятилетие бум в области искусственного интеллекта. Вместе с тем по настоящее время не ясно, что представляет собой интеллект естественный.

Существование различий между людьми по умению решать задачи одного класса неоспоримо. Это позволило выдвинуть предположение о существовании в человеке особых ментальных структур, умственных способностей, отраженных в понятии «интеллект». Именно на этом предположении был основан самый известный классический тест для оценки коэффициента интеллекта (*IQ, intelligence quotient*) – шкала Д. Векслера, которая по настоящее время служит эталоном для оценки человеческого ума в его рациональных и культурно-лингвистических формах.

В современной интерпретации, предложенной в работах М.А. Холодной, интеллект – это форма организации когнитивного опыта человека, представленного в виде «накопленных» в ходе онтогенеза понятийных психических структур, степень сформированности которых определяет структурные характеристики субъективного пространства

интеллектуального отражения. Основное назначение интеллекта по М.А. Холодной – построение особого рода репрезентаций происходящего, связанных с воспроизводством объективного знания о мире [6]. Отметим, что эта задача решается посредством создания и использования когнитивных инструментов, которые должны быть достаточно удобны и эффективны для пользователя.

В классических, главным образом, факторных моделях интеллекта рассматривается его адаптивная функция по отношению к окружающей и внутренней средам. Определение интеллекта как некоторой способности, обуславливающей успешность адаптации индивида к новым условиям, является наиболее общим местом в моделях, использующих методологию классической научной рациональности [7].

В классическом подходе среда рассматривается главным образом как отдельный фактор, влияющий на формирование и проявления интеллекта, а не как его неотъемлемая часть. Эта отделенность интеллекта от среды создает иллюзии его независимости и локализации в субстрате организма человека.

При изучении интеллекта никто не обращает внимания на его активный, организующий характер при формулировании задач, их оптимизацию под когнитивные средства и возможности, имеющиеся у конкретного обладателя интеллекта. Естественный интеллект активен и избирателен по своей природе. Он упрощает среду деятельности до уровня, позволяющего ее операционализировать и тем самым активно преобразовывать в нужном направлении.

В модели интеллекта Д. Гилфорда показаны возможности человека по универсальному структурированию среды. Выделено 120 интеллектуальных способностей (позднее их число было увеличено до 150), каждая из которых связана с определенным классом задач, представленных в координатах: содержание, операция, продукт.

Р. Мейли выдвинул гипотезу о том, что структуру

интеллекта составляют четыре фактора: доступной сложности, пластичности, целостности, беглости. На основе проведенных исследований предложена «компонентная модель» интеллекта, в которой факторы трактовались как условия (компоненты) индивидуальных различий в выполнении разных интеллектуальных актов. Они могут относиться как к индивиду, так и к окружающей среде. Следовательно, структура интеллекта, включающая эти факторы, включает в себя структуру взаимодействия индивида со средой.

В теории множественного интеллекта Говарда Гарднера выделено девять типов интеллекта: лингвистический, логико-математический, визуально-пространственный, телесно-кинестетический, музыкальный, натуралистический, межличностный, внутриличностный и экзистенциальный. Все виды интеллектов по Гарднеру равноценны, каждый представляет собой особый способ взаимодействия человека с окружающей действительностью. Очевидно, что каждый интеллект в данной модели эффективен только по отношению к специализированной среде, отражающей соответствующие формы культуры.

К интеллектным симбионтам относятся различные виды интеграции интеллектуальной и когнитивной системы человека с различного рода физическими сущностями, позволяющими повысить эффективность возникающей системы при решении интеллектуальных задач. Такими объединениями могут быть системы гибридного и диффузного интеллектов, применение которых позволяет обеспечить эффективность выполнения сложных видов деятельности.

Гибридные модели связаны с разработкой практических принципов взаимной адаптации человека с техникой и условиями труда, созданием теории и методов синтеза систем адаптивного взаимодействия людей между собой и с интеллектуальной системой.

Термин «гибридный интеллект» (ГИ) был впервые

введён В.Ф. Вендой в 1975 году на конференции по семантическим вопросам искусственного интеллекта, а основы теории систем гибридного интеллекта изложены в изданном в 1977 году сборнике по инженерной психологии [1].

Теория систем гибридного интеллекта построена на основе общих для всех видов систем законов. В качестве таковых В.Ф. Вендой предложены законы взаимной адаптации и трансформации систем. Взаимная адаптация акцентирует внимание на изменениях, которые претерпевают объекты, становясь компонентами системы, на закономерностях этих изменений в ходе становления, развития, существования, трансформации структур системы. Взаимная адаптация человека и компьютерной системы направлена на максимальное раскрытие индивидуальных способностей, компенсацию психофизиологических недостатков, учёт интересов лица принимающего решения. Этот процесс, по мнению В.Ф. Венды, также ведёт к наиболее полному использованию возможностей вычислительной техники [4].

Принципиальное отличие методологии систем гибридного интеллекта от традиционной методологии инженерной психологии состоит в том, что вместо анализа вариантов и попытки выбрать из них оптимальный проводится синтез разных вариантов решений интеллектного объединения скрытых и непосредственных участников решения. Системы гибридного интеллекта рассматриваются как комбинированные системы, интегрально включающие в себя искусственный и естественный интеллекты.

Вместе с тем следует отметить, что технологически эта теория реализует анализа кривых обучения в процессе адаптации [4]. Процесс обучения представлен в виде комплекса кривых параболического типа с возрастающими значениями максимума качества деятельности по мере освоения более совершенных стратегий, на основе которых строятся немонотонные (с промежуточными плато и горбами) кривые обучаемости. Теория основана на критике

монотонных кривых обучения, что справедливо, но она не имеет под собой психологического обоснования и содержания, будучи механистичной, инженерной интерпретацией работы живых и социальных систем. Ключевое понятие теории «стратегия» неопределенно и связано лишь с наблюдаемым поведением и показателями качества деятельности, которые в свою очередь зависят от субъективной сложности задач. В силу этого теория не получила распространения в отечественной психологии и инженерной практики.

Интеллект и организованность среды

Интеллект связан с миром задач, различающихся между собой сложностью решения, и может интерпретироваться с позиций концепций сложности. Выделяют «онтологическую» и «семиотическую» простоту/сложность. Под первой понимают сложность материального, физического мира, а под второй – оценку знаковых систем. Понятие интеллекта в большей мере связано с семиотической сложностью, хотя мир задач обусловлен и онтологической сложностью. Отношения между этими видами сложности и определяют средово-системный континуум, в котором порождается понятие «интеллект».

Таким образом, традиционное понимание интеллекта – как суммы локально принадлежащих когнитивному аппарату человека способностей (инструментальных возможностей искусственной системы в случае искусственного интеллекта) – должно быть заменено на интегральные свойства, обусловленные средой и действующей (живущей) в ней системой. Интеллект определяется в значительной мере степенью освоенности конкретной среды системой, действующей в ней, степенью воплощенности в нее. Отметим, что естественный интеллект связан с механизмами управления и целеполагания человека и говорить об интеллекте как независимой сущности нельзя. Вместе с тем близкая категория - «разум» менее инструментальная, чем

интеллект.

Большинство исследователей сложного считают, что представления о сложном могут быть сформулированы в концептах: множественности; динамического разнообразия; нелинейности; неравномерности; сложности самоорганизующихся систем [8]. Эти же категории применимы и к понятию «интеллект».

Множественность рассматривается как многокомпонентность. Она относится к описанию сложных систем непрерывно эволюционирующих и изменяющихся. Сложная система в этой парадигме предстает как процесс бесчисленного усложнения ее сущностей, возникновения новых элементов как элементов для новых творений. Интеллект может быть представлен как свойство, воплощенное в сложную динамическую систему, позволяющее достигать результат, определяемый условиями задачи.

Множественность в концепциях динамического разнообразия дополняется качественной характеристикой – разнообразием. Разнообразие связано с асинхронным существованием в среде динамически существующих и сосуществующих систем и их распадающихся элементов, которые могут образовывать, в свою очередь, новые системные сущности.

Концепт неравномерности отражает принципиальную неравномерность распределения в пространстве одновременно существующих различных форм материи (энергии, вещества), ведущую к локальной самоорганизации и возникновению новых систем. Постулируется холистический характер мира, который разделяется на элементы только в процессе работы механизмов человеческого интеллекта и сознания.

Концепция сложности самоорганизующихся сред и систем отражает непрерывную динамику мира во всех его проявлениях, а интеллект в ней является эмерджентным свойством сложной системы, позволяющим последней эффективно решать задачи активного (формирование искус-

ственной среды, созидающая деятельность) и пассивного (адаптация, приспособление к среде) существования в мире [9].

Организованность среды определяет возможности по ее освоению интеллектными системами. В случае низкого уровня организованности можно говорить о среде деятельности, а в высокоорганизованных средах – о содействующих системах. Интеллектуальные системы призваны снизить степень сложности среды, организовать ее в доступной для оператора системы форме.

Интеллектные объединения (симбионты)

Сложные (наделенные интеллектом) системы являются системами операционально закрытыми. Система одновременно является открытой и замкнутой по отношению к окружающей среде. Операциональная замкнутость означает селективность системы, наличие границы, упорядочивающей отношения системы со средой и окружающими системами. Система и среда проявляют взаимную активность. Среда меняет систему, но и система активно видоизменяет окружающую среду, вступая в коэволюцию с нею.

Сложность является возникающим и исчезающим феноменом, циклически порождаемым в циклах самоорганизации. Многие ее аспекты перекликаются с понятием интеллекта. Можно предположить, что интеллектуальные функции отражают некоторые общие принципы самоорганизации аутопоэтических систем, в частности, их способность активно изменять сложность среды [9]. Отсюда следует, что все аутопоэтические системы, в том числе и образуемые социальными коммуникациями, обладают воплощенными в них интеллектуальными способностями. Следует заметить, что все системы, проявляющие интеллектуальное поведение (в том числе и технические) непрерывно действуют в циклах «редукции-усложнения» своего когнитивного содержания [10]. Данный механизм позволяет сис-

теме различать внутреннее и внешнее, формировать динамическую границу и выбирать требуемый уровень сложности задачи. Это связано также с попыткой системы найти наиболее эффективное в данных ресурсных ограничениях решение.

В таблице 1 приведены основные свойства интеллектных образований, возникающих в эргатических системах и техногенных средах использующих технологии искусственного интеллекта. Акценты сделаны на отношениях элементов возникающих систем между собой, месте центра активности объединенной системы и отношениях объединяющихся систем со средой деятельности. Подчеркивается определенная независимость границ объединяющихся систем. Отметим, что в интеллектных симбиотических системах функции редукции сложности чаще всего выполняются человеком. При этом качество решения зависит от качества информации привлекаемой из внутренней и внешней сред эргатической системы (среды).

Дальнейшее развитие процессов интеграции искусственного и естественного интеллектов ведет к появлению новой формы интеграции – диффузного интеллекта, который является формой усиления когнитивной системы человека за счет программно-технических элементов сложноорганизованной информационно-коммуникативной среды. Данная форма интеграции зависит от качества систем интерфейса объединяющих когнитивные элементы человека и технической системы наделенной искусственным интеллектом.

Таблица 1 - Свойства интеллектных симбионтов

Вид интеллектного объединения	Связи между компонентами	Источник активности и механизм управления	Отношения со средой деятельности, свойства границ
Естественный интеллект	Самоорганизация, аутопоэтическая система	Субъективная реальность, сознание, эгосистема человека	Активное преобразование мира, границы динамично изменяемые
Искусственный интеллект	Фиксированная или переменная программно-аппаратная структура	Программа, алгоритм в структурированной или структурируемой среде	Реализация алгоритма, ситуативное управление, границы фиксированные
Гибридный интеллект	Симбиоз, адаптация организованного и аутопоэтического компонентов к среде, объединения на макроуровне с приоритетом сознания	Человек в структурированной среде	Взаимная адаптация естественного и искусственного интеллектов, границы переменные
Диффузный интеллект	Селективные связи на всех уровнях аутопоэтически организованной и организуемой среды и человека	Возникает в организованной среде действительности, физической реальности и программной среде	Синергетическое объединение субъективной реальности, границы формируются под задачу

Интеллект эргатических систем с искусственным интеллектом

Любая интеллектуальная эргатическая система представляет собой, в известной мере, автономную систему, осуществляющую с помощью оператора деятельность по достижению заданной цели в конкретной среде. Чем больше функций во взаимодействиях со средой без присутствия оператора может выполнять система, тем более сложной и интеллектуальной она кажется наблюдателю.

Интеллект искусственной технической системы возникает и проявляется в условиях организованной сложности в отличие от живых систем, активно организующих среду. Система искусственного интеллекта эргатической системы является системой организованной сложности, в которой взаимодействие элементов и функциональных областей строятся на неслучайных отношениях, заданных проектировщиком в некоторых контекстах, которые непрерывно меняются, следуя логике эволюции задачи, вызываемой оператором.

Функциональная сложность эргатической системы может быть оценена по аналогии с когнитивной сложностью человека по количеству и типам решаемых задач. Ее сложность (и интеллектуальность) может определяться в привязке к конкретным решаемым задачам, к сложности программных средств, используемых в управляющем компьютере системы, с объемом и степенью дифференцированности контактов системы со средой и оператором. В качестве критерия сложности (интеллектуальности) может использоваться длина выполняемого алгоритма и количество циклов, позволяющих его реализовать.

Перспективным методом определения интеллектуальности эргатической системы является метод, учитывающий свойства культурного поля проектного коллектива, участвующего при создании системы. Чем шире спектр специалистов и более высокая междисциплинарная нагрузка, тем система сложнее и интеллектуальнее.

Появление эргатических систем и сред с искусственным интеллектом требует рассматривать вопросы взаимодействия естественного интеллекта оператора системы с искусственным интеллектом, воплощенным в эргатической системе / среде.

Можно дать ряд общих определений интеллекта и интеллектных симбионтов, действующих в эргатической системе и возникающих в процессе объединения искусственного и естественного интеллектов и среды деятельно-

сти.

1. Интеллект есть форма активной самоорганизации сложной системы, вовлекающая погруженного в среду пользователя в созидающие изменения.

2. Интеллект связан со средой как механизм ее организации, обеспечивающий процессы самоорганизации системы им наделенной.

3. Интеллект распределен в континууме «система-среда» и воплощен в циклах самоорганизации системы, действующей в среде.

4. Естественный интеллект представляет собой организующую сложность в организуемой среде, а искусственный интеллект – организованную сложность в организованной среде.

5. Гибридный и диффузный интеллекты представляет собой симбионты, включающие организующую и организованную сложность систем в их синергетическом и симбиотическом взаимодействиях как инструмент достижения цели актором в организованной и организуемой средах.

6. Интеллект отражает результаты селекции и применения самоорганизующейся системой эффективных способов достижения цели в организованной среде.

Заключение

Интеллектные симбионты являются проявлением различных форм межсистемных объединений различной материальной природы, в которых проявляется активная форма преобразования и структурирования среды.

Гибридные формы интеллектных симбионтов представляют собой объединения, не нарушающие системную целостность составляющих их симбионтов. При этом совместное функционирование симбионтов порождает новые, полезные для гибридной системы, интеллектуальные свойства.

При проектировании сложных эргатических систем важно правильно выбрать методы и технологии их интеллектуализации, которые в основном зависят от среды деятельности. Попытка полной автономизации и автоматизации системы связана со степенью организованности среды. При работе в условиях организованных сред целесообразно использование методов распределенного искусственного интеллекта, а в слабоструктурированных средах – включение и усиление возможностей естественного интеллекта.

При создании сложных эргатических систем необходимо учитывать эмерджентные свойства, возникающие вследствие сложной организации среды. Это эффекты интеллектуализации, кооперативных и гибридных форм интеллектов и их симбионтов. Включение в сложные техногенные среды человека также связано с эффектами модификации его личности и когнитивных систем, что ведет к возникновению техно-психических симбиозов. Например, память человека заменяется памятью технической системы, функции внимания и поиска информации – работой поисковых систем и т.д.

Материалы настоящего исследования могут быть использованы при создании систем интерфейсов сложных эргатических систем, включающих симбиотические объединения аутопоэтического типа, возникающие в распределенных эргатических системах с социальными формами управления и взаимодействия с операторами.

Список литературы

1. Инженерная психология: Теория, методология, практическое применение. Отв. ред. Б.Ф. Ломов, В.Ф. Рубахин, В.Ф. Венда. – М., 1977. – 304 с.
2. Сергеев С.Ф. Методология проектирования тренажеров с иммерсивными обучающими средами // Научно-технический вестник СПбГУ ИТМО. 2011. № 1 (71). – С. 109-114.
3. Сергеев С.Ф. Методология создания мехатронных сис-

тем с искусственным интеллектом // Известия ТулГУ. Тех-Технические науки: В 3 ч. Ч. 1. – Тула: Изд-во ТулГУ, 2011. – Вып. 5. – С. 245-249.
4. Венда В.Ф. Системы гибридного интеллекта: Эволюция, психология, информатика / В.Ф. Венда. – М.: Машиностроение, 1990. – 448 с.
5. Сергеев С.Ф. Психологические основания проблемы искусственного интеллекта // Мехатроника, автоматизация, управление. 2011. № 7. – С. 2-6.
6. Холодная М.А. Когнитивные стили. О природе индивидуального ума. – М.: ПЕР СЭ, 2002. – 304 с.
7. Степин В.С. Классика, неклассика, посттнеклассика: критерии различения // Постнеклассика: философия, наука, культура. – СПб: Изд. дом Міръ, 2009. – С. 249-295.
8. Князева Е.Н. Темпоральная архитектура сложности // Синергетическая парадигма. Синергетика инновационной сложности. – М.: Прогресс-Традиция, 2011. – С. 66-86.
9. Сергеев С.Ф. Проблема сложности в эргатических системах // Материалы конференции «Управление в технических, эргатических, организационных и сетевых системах» (УТЭОСС-2012). – СПб.: ГНЦ «Концерн «ЦНИИ «Электроприбор», 2012. – С. 802-805.
10. Сергеев С.Ф. Роль механизма редукции в обучении и образовании // Философия образования. 2013. № 1(46). – С. 198-205.

УДК 519.7:681.31
И.Ф. Чебурахин, д-р техн. наук, проф.
МАТИ - РГТУ им. К.Э. Циолковского, г. Москва, Россия
cybernetics@mati.ru

ОБРАБОТКА ДАННЫХ И УПРАВЛЕНИЕ В ДИСКРЕТНЫХ СИСТЕМАХ НА ОСНОВЕ ФУНКЦИОНАЛЬНО-ЛОГИЧЕСКИХ УРАВНЕНИЙ

Современные достижения в области науки и техники основываются на использовании прямо и косвенным образом различных информационно-управляющих систем. Данные системы находят широкое распространение во всех областях человеческой деятельности, включая промышленность, энергетику, транспорт и связь. Ключевую роль в данных системах играют устройства логического управления и обработки данных. В докладе обсуждаются разработанные автором методы реализации логических алгоритмов (булевых функций) на основе структурно-функциональной параллельной (логической) декомпозиции [1-4], позволяющие, наряду с синтезом схем, получать соответствующие им значения различных показателей сложности в разных базисах.

Основные понятия и обозначения

Введем обозначения: $X = \{x_1, \ldots, x_n\}$ - множество булевых переменных; $f^{(n)}$ - булева функция (БФ) над множеством X. Под базисом G понимаем конечную функционально полную систему БФ (или соответствующих функциональных элементов - ФЭ). Рассматриваем базисы $G_1 = \{\&, \vee, \ \}$ и $G_3 = \{\&, \oplus, 0, 1\}$ для всех БФ и - $G_2 = \{\&, \vee\}$ для монотонных функций. Считаем, что функция $f^{(n)} \in P_2$ задается формулой $F^{(n)}$ (или F) в базисе G.

В качестве меры сложности представления функции f формулой F или схемой S из ФЭ определяем соответст-

вующие показатели (дискретные функционалы): $L_Б(f, G)$ – суммарное число вхождений символов переменных (букв) в формулу F, реализующую функцию f в базисе G; $L_F(f, G)$ - число подформул (базисных) в F; $L_S(f, G)$ - число ФЭ в схеме S, реализующей f.

Пример функционального уравнения (ФУ):

$$f^{(n)} = h^{(3)}(F^{(n-1)}, f^{(n-1)}, x_n), \quad (1)$$

где $h^{(3)}$ - трехместная функция рекурсии, $f^{(n)}$ - исследуемая функция.

Рассматриваемые методы синтеза устройств логического управления и обработки информации

Методы синтеза устройств логического управления и обработки информации позволяют минимизировать значения оценок сложности и/или минимизировать требуемые вычислительные ресурсы. К данному классу методов следует отнести следующие методы (краткая их функциональная характеристика дана в работах [1-4]):

♦ **Синтез схемы S на основе С-метода.** Данный (упрощенный, «скорый») метод позволяет выводить оценки показателей сложности $L_Б$, L_F и L_S полинома Жегалкина $F^{(n)} = K_1 \oplus \ldots \oplus K_m$, на основе его строения $\boldsymbol{r} = (r_1, \ldots, r_m)$. Метод дает грубые оценки, но они получаются просто, быстро, с минимальной трудоемкостью.

♦ **Синтез схемы S на основе метода суперпозиции.** Здесь формула $F_{\text{с.ф.}}^{(n)}$ получается в базисе Жегалкина с помощью структурно-функциональной параллельной (логической) декомпозиции (СФПаД) и хранится в специальной таблице. По полученной таблице каждая базисная функция заменяется на соответствующий ФЭ с сохранением связей.

♦ **Синтез схемы *S* методом структуризации.** Здесь в полиноме Жегалкина $F^{(n)}$ выделяется однородная структура на основе повторяющегося фрагмента, начальная и завершающая структуры, а также определяются показатели сложности L_F (и L_S) всех трех структур. Их сумма позволяет априори находить и минимизировать показатели сложности L_F (и L_S) представления БФ $f^{(n)}$ в классах формул $F^{(n)}$ и схем *S*.

♦ **Синтез схемы *S* на основе метода ветвления.** Данный метод предполагает для каждых равных подформул реализацию одной из них на соответствующем ФЭ и подсоединения его выхода к определенным входам адекватных ФЭ схемы**.** Для получения такой возможности над формулой выполняются требуемые эквивалентные преобразования.

Представление БФ в классах формул и схем на основе СФПаД: эквивалентные преобразования формул и оценка сложности

В задачах дискретного логического управления и обработки информации все большее распространение находят интеллектуальные технологии. При этом весьма перспективным является использование т.н. агентных технологий [5-6]. Так, например, при использовании агентных технологий в беспроводных компьютерных сетях для решения задач дискретного логического управления сложными техническими системами (коллективом объектов) алгоритм управления, в общем случае, есть система БФ [6]. Здесь в качестве агента выступает программа, вычисляющая определенную булеву формулу. Для закона управления есть полная система программ (агентов), вычисляющих соответствующие формулы $F^{(n)}$. С головного PC-host по-

ступают начальные данные на управляющий РС, они рабатываются непрямыми методами, позволяющими локально для исходных данных минимизировать число вычислительных операций, и результат от всех агентов, поступает на головной PC-host для принятия решения по управлению.

Непрямые методы понимаем как эффективную обработку данных: вычисления выполняются не по формуле (полиному Жегалкина, ДНФ и др.), а по другой модели - информационной логико-математической, основанной на строении формулы. Каждой ситуации соответствует определенный набор значений переменных. При аппаратной реализации логических алгоритмов применяют распараллеливание их на внутрикристальном уровне.

Функции из классов “&”, ”∨”, “⊕” или “~” обладают одинаковыми алгебраическими свойствами. Поэтому, какую структуру выводим для одной из них, будет иметь место для каждой из них. При представлении функции $f^{(n)}$, конечное $n \geq 2$, формулой над $\{g^{(m)} \mid$ конечное $m \geq 2\}$, на основе СФПаД получаем минимальное значение показателя сложности L_F представления в классах формул и логических схем без ветвления:

$$L_F(F^{(n)}, g^{(m)}) = L_S(F^{(n)}, g^{(m)}) = \lceil (n-1)/(m-1) \rceil.$$

Для класса симметрических БФ получаем элементарные симметрические полиномы (ЭСП) Жегалкина $F_i^{(n)}$, где n - конечное число переменных, нижний индекс i - степень полинома, $i = \overline{1, n}$.

Так, для полиномов $F_1^{(n)}$ (класс функций - “⊕”) и $F_n^{(n)}$ (класс “&”) показатели сложности L_F его минимального представления изучены ранее в классе формул:

$$L_F^{\min}\left(F_1^{(n)}, G_3\right) = L_F^{\min}\left(F_n^{(n)}, G_3\right) = n-1, \qquad (2)$$

для полиномов $F_i^{(n)}$, степени $i = 2$ или 3, верхние оценки следующие:

$$L_F\left(F_2^{(n)}, G_3\right) = (n^2 + n - 4)/2\,,\ L_F\left(F_3^{(n)}, G_3\right) = (n^3 - n - 12)/6\,. \qquad (3)$$

Это же будут и верхние оценки соответствующих схем S (без ветвления):

$$L_S\left(F_2^{(n)}, G_3\right) = (n^2 + n - 4)/2\,,\ L_S\left(F_3^{(n)}, G_3\right) = (n^3 - n - 12)/6\,. \qquad (4)$$

На основе анализа оценок (2)-(4) сложности L_S представления произвольного ЭСП Жегалкина (в классе схем с ветвлением) получим их обобщение по двум показателям: по числу переменных n и степени i полинома Жегалкина, т.е. оценку (5), вычисляемую за чуть больше 10 арифметических операций:

$$L_S(F_i^{(n)}, G_3) = (n-1) + (i-1)(2n-4) - (2i^2-6i+4),\ i = \overline{1,n}\,. \qquad (5)$$

Аналогичный подход к синтезу логических схем имеет место также для класса монотонных БФ в базисе G_2.

Применим метод ФУ для представления произвольной БФ $f^{(n)}$ формулой $F_{\text{с.ф.}}^{(n)}$ в базисе G_3 и получения при этом верхней оценки сложности L_F

$$L_F(F^{(n)}, G_3) = (5/4)\cdot 2^n - 2. \qquad (6)$$

Оценку (6) сложности L_F представления произвольной БФ $f^{(n)}$ в базисе G_3 уточним для отдельных классов функций из P_2. Пусть функция $f^{(2)}$ задается полиномом

$$F^{(2)} = x_1x_2 \oplus x_1 \oplus x_2 \oplus 1,$$

преобразуем его эквивалентным способом к виду

$$F^{(2)} = (x_1 \oplus 1)(x_2 \oplus 1),$$

причем значение показателя сложности L_F минимально $L_F^{\min}\left(F^{(2)},\ G_3\right) = 3$.

Индукцией доказываем (для конечного натурального *n* выделяем класс $f^{(n)}$):

$$F^{(n)} = x_1 \cdot x_2 \cdot \ldots \cdot x_n \oplus x_1 \cdot x_2 \cdot \ldots \cdot x_{n-1} \oplus x_2 \cdot \ldots \cdot x_n \oplus \ldots \oplus x_1 \oplus x_2 \oplus \ldots \oplus x_n \oplus 1 =$$ (7)

$$= (x_1 \oplus 1) \cdot (x_2 \oplus 1) \cdot \ldots \cdot (x_{n-1} \oplus 1) \cdot (x_n \oplus 1),$$ (8)

$$L_F^{\min}\left(F^{(n)}, G_3\right) = 2n - 1.$$ (9)

Преобразуем (8) в базисе G_1, применяя правило де Моргана, получаем

$$F^{(n)} = \overline{x_1 \vee x_2 \vee \ldots \vee x_n},$$ (10)

$$L_F^{\min}\left(F^{(n)}, G_1\right) = n.$$ (11)

Сравнивая оценки (9) и (11) сложности представления функции $f^{(n)}$ в классе формул, получаем

$$L_F^{\min}\left(F^{(n)}, G_3\right) \ / \ L_F^{\min}\left(F^{(n)}, G_1\right) = (2n-1)/n = 2 - 1/n.$$ (12)

Примем следующие оценки:

$$\boldsymbol{L_{\min} = L_F^{\min}\left(F^{(n)}, G_3\right) = 2n - 1,\quad L_{\text{асим}} = L_F^{\text{асим}}\left(F^{(n)}, G_3\right) = 2^n / n,}$$

$$\boldsymbol{L_{\text{верх}} = L_F^{\text{верх}}\left(F^{(n)}, G_3\right) = (5 \cdot 2^n / 4) - 2.}$$

Вывод: Как следует из (12), функции $f^{(n)}$ одного класса (подмножества P_2), задаваемые полиномами Жегалкина строения типа (7), при представлении в базисах G_1 и G_3 имеют один порядок роста.

Итак, вычисления показывают, что при $2<n<7$ оценка $L_{\text{асим}}$ занижена относительно - $L_{\min}$ ($L_{\min} > L_{\text{асим}}$), но при

n>6 оценка $L_{асим}$ завышена относительно - L_{min} ($L_{min} < L_{асим}$) и расхождение продолжает увеличиваться при возрастании n.

Алгоритмизация определения верхней оценки Сложности представления БФ в классе формул в базисе G_3

Исследуем показатель сложности $L_Б$. Разбиваем эффективным способом полином Жегалкина $F^{(n)}$ на два полинома меньшей сложности, затем каждый из них разбиваем опять на два и так далее. Получаемое в результате дерево решений будет учитывать особенности задачи (сложные или неопределенные параметры). На некотором шаге вычисления по определенной цепочке могут заканчиваться, а по другим - продолжаться. Именно это делает алгоритм приспосабливающимся к исходным данным задачи.

Пусть $F^{(n)} = K_1 \oplus \ldots \oplus K_i \oplus \ldots \oplus K_m$, $i = \overline{1,m}$, где K_i - монотонная элементарная конъюнкция (ЭК) ранга r_i, причем $r_1 \geq \cdots \geq r_i \geq \cdots \geq r_m$, $i = \overline{1,m}$. Такой полином Жегалкина называется каноническим, его ЭК расположены в соответствии с порядком "≥", определенным для их рангов. Вектор $\boldsymbol{r} = (r_1, \ldots, r_i, \ldots, r_m)$ задает строение полинома Жегалкина $F^{(n)}$. При этом определяем еще вектор $\boldsymbol{p} = (p_1, \ldots, p_j, \ldots p_n)$ повторяемости переменных из множества X в формуле $F^{(n)}$, т.е. переменная x_j, $j = \overline{1,n}$, повторяется в формуле $F^{(n)}$ число p_j раз.

Опишем в общем виде методику получения верхней оценки сложности $L_Б(F^{(n)}, G_3)$ с помощью СФПаД:

$$F^{(n)} = \left((x_{i1} \cdot F_0^{(n-1)}) \oplus F_1^{(n-1)} \right),\ i1 = \overline{1, n-1}, \qquad (13)$$

где нижние индексы 0 и 1 – номера соответствующих оста-

точных функций-полиномов Жегалкина $F_0^{(n-1)}$ и $F_1^{(n-1)}$. Эти функции формально зависят от одного числа $n-1$ переменных множества $X' = X \setminus \{x_{i1}\}$.

Найдем максимум

$$p_j = \max(p_1, \dots, p_j, \dots, p_n), j = \overline{1,n}.$$

Применим (13) для представления полинома Жегалкина $F^{(n)}$:

$$F^{(n)} = x_{i1} \cdot K_1 \oplus x_{i2} \cdot K_2 \oplus \dots \oplus x_{ij} \cdot K_j, \qquad (14)$$

где полиномы K_j, $j = \overline{1,n-1}$, рассматриваются на множестве $X \setminus \{x_{i1}, x_{i2}, \dots, x_{ij}\}$. К полученным полиномам K_j применяем указанную технологию.

Введем оператор $L_Б$ (L_F, L_S и возможно другие, имеющие свои функциональные особенности) для чтения рекуррентных соотношений типа (13), как в математической логике. Тогда получаем следующее функциональное уравнение показателя сложности (ФУ$_{ПС}$) $L_Б$, для которого выводим ряд оценок

$$L_Б(F^{(n)}) \geq L_Б\left((x_j \cdot F_0^{(n-1)}) \oplus F_1^{(n-1)}\right) = L_Б\left((x_j \cdot F_0^{(n-1)}) \oplus L_Б(F_1^{(n-1)})\right) =$$

$$= L_Б(F_0^{(n-1)}) + L_Б(F_1^{(n-1)}) + 1 =$$

$$L_Б(F_0^{(n-1)} \oplus F_1^{(n-1)}) + 1 = L_Б(F^{(n-1)}) + 1, \qquad (15)$$

где

$$F^{(n-1)} = F_0^{(n-1)} \oplus F_1^{(n-1)} \text{ и } L_Б(F^{(n)}) > L_Б(F^{(n-1)}).$$

При этом соответствующий полином применяется для синтеза суперпозиционной формулы $F_{с.ф.}^{(n)}$ и получения верхней оценки сложности $L_Б$, эффективно уменьшая вычислительную сложность.

Для полинома Жегалкина $F^{(n)}$ составляется таблица 1 и подсчитываются числовые характеристики: векторы рангов $\boldsymbol{r}$ и повторяемости $\boldsymbol{p}$ переменных, а также - исходная сложность $L_Б(F^{(n)})$ полинома. Представление полинома Жегалкина $F^{(n)}$ при помощи таблицы 1 помогает проведению конструктивных преобразований, при необходимости параллельно дублируемых аналитически.

Таблица 1

$\boldsymbol{x}$ / K	x_1	…	x_j	…	x_n	$\boldsymbol{r}$
K_1	…	…	…	…	…	r_1
…	…	…	…	…	…	…
K_i	…	…	$x_{i,j}=1$	…	…	r_i
…	…	…	…	…	…	…
K_m	…	…	…	…	…	r_m
$\boldsymbol{p}$	p_1	…	p_j	…	p_n	$L_Б$ = L

Алгоритм

N - число циклов; $L_Б$=L - текущая сложность полинома по числу букв.

Шаг 1. Для заданного полинома Жегалкина $F^{(n)}$ заполнить табл. 1.

Присвоить N := 0. Элемент таблицы $x_{i,j}=1$, где i - номер строки (ЭК K_i), j - номер столбца для переменной x_j.

Шаг 2. Подсчитать: векторы рангов r и повторяемости p переменных; $L := \Sigma_{i=1}^{m} r_i$.

Шаг 3. По таблице 1 определить $p_j = \max(p_1, \dots, p_j, \dots p_n)$, $j = \overline{1,n}$ и выбрать

x_j, $j = \overline{1,n}$. N := N + 1; $L := L - p_j + 1$.

Шаг 4. Из полинома Жегалкина $F^{(n)}$ сформировать на основании табл. 1 полиномы $F_0^{(n-1)}$ и $F_1^{(n-1)}$:

- остаточные функции в (13) не содержат переменную x_j, но в требуемую суперпозиционную формулу $F_{\text{с.ф.}}^{(n)}$ включаем ее множителем;

- выбираем число p_j ЭК K_i, содержащие переменную x_j; их ранги r_i уменьшаем на единицу; перемещаем эти строки в верхнюю часть табл. 1 так, чтобы сохранить порядок "≥" для рангов r_i; получаем полином $F_0^{(n-1)}$.

- ЭК K_i, не содержащие переменную x_j, располагаем в нижних строках табл. 1, сохраняя для рангов r_i ЭК K_i порядок "≥"; получаем полином $F_1^{(n-1)}$.

Получаем рекуррентное соотношение (13), для которого вычислительным методом получено значение показателя сложности $L_{\text{Б}} = L$.

Шаг 5. На основании принципа "для каждой остаточной функции выбирается своя максимально повторяющаяся переменная" преобразовать полученные полиномы $F_0^{(n-1)}$ и $F_1^{(n-1)}$ для дальнейшей их минимизации:

- на основе таблицы 1 получаем новый вектор строения $\boldsymbol{r'}$, а также вектор повторяемости переменных $\boldsymbol{p}$;

- определяем $p_{t,0} = \max(p_1, \dots, p_t, \dots p_{n-1})$ и $p_{t,1} = \max(p_1, \dots, p_t, \dots p_{n-1})'$, $t = \overline{1, n-1}$, и выбираем соответствующие переменные $x_{t,0}$ и $x_{t,1}$, $t = \overline{1, n-1}$.

Далее, аналогично предыдущему, выбираем требуемые $p_{t,0}$ (или $p_{t,1}$) ЭК $K_{i,0}$ ($K_{i,1}$), которые содержат соответствующую переменную $x_{t,0}$ (или $x_{t,1}$) вынести по-

очередно за скобки переменную x_t, соответственно корректируя таблицу 1, т.е. выполняем эквивалентные преобразования подобно тому, как делалось на шаге 4 при получении (13), при этом $L := L - p_{t,0} - p_{t,1} + 2$.

В результате, получаем рекуррентные соотношения типа (13),

$$F_0^{(n-1)} = \left((x_{t,0} \cdot F_{00}^{(n-2)}) \oplus F_{01}^{(n-2)} \right),$$

$$F_1^{(n-1)} = \left((x_{t,1} \cdot F_{10}^{(n-2)}) \oplus F_{11}^{(n-2)} \right),$$

для которых для данных шагов получены показатели сложности типа (14).

$$L_Б\left(F_0^{(n-1)}\right) = L_Б\left(F_{00}^{(n-2)}\right) + L_Б\left(F_{01}^{(n-2)}\right) + 1,$$

$$L_Б\left(F_1^{(n-1)}\right) = L_Б\left(F_{10}^{(n-2)}\right) + L_Б\left(F_{11}^{(n-2)}\right) + 1.$$

Следовательно,

$$L_Б\left(F^{(n)}\right) = L_Б\left(F_0^{(n-1)}\right) + L_Б\left(F_1^{(n-1)}\right) + 1 =$$

$$= L_Б\left(F_{00}^{(n-2)}\right) + L_Б\left(F_{01}^{(n-2)}\right) + L_Б\left(F_{10}^{(n-2)}\right) + L_Б\left(F_{11}^{(n-2)}\right) + 3.$$

Далее, аналогично, получаем $F^{(n-2)}$ и т.д., пока в таб. 1 не останутся две переменные, с которыми завершится построение формулы $F_{\text{с.ф.}}^{(n)}$, иначе перейти к шагу 4. При этом, если в формуле выходим на базисные функции (“$x_1 \& x_2$” или “$x_1 \oplus x_2$”, или большей местности таких функций, то для повышения практической значимости алгоритма применить параллельную декомпозицию.

Конец работы алгоритма.

Таким образом, на основе алгоритма конструктивно получаем последовательность полиномов Жегалкина $F^{(n)}$,

сложность $L_Б$ которых уменьшается:

$$L_Б(F^{(n)}, G_3) \geq L_Б(F^{(n-1)}, G_3) \geq L_Б(F^{(n-2)}, G_3) \geq \ldots \geq L_Б(F^{(2)}, G_3),$$

а также - суперпозиционную формулу $F_{\text{с.ф.}}^{(n)}$ для построения схемы *S* с оценкой сложности L_S.

Теперь применяем метод синтеза схемы *S* на основе суперпозиционной формулы $F_{\text{с.ф.}}^{(n)}$, получаемой в базисе Жегалкина с помощью структурно-функциональной (параллельной) декомпозиции. По формуле $F_{\text{с.ф.}}^{(n)}$ каждая базисная функция заменяется на соответствующий ФЭ с сохранением связей, при этом создается требуемая схема *S* с оценкой сложности L_S.

Заключение

Разработанные методы реализации логических алгоритмов (булевых функций) на основе структурно-функциональной параллельной (логической) декомпозиции позволяют наряду с синтезом схемы, получать соответствующие значения различных показателей сложности в разных базисах. Применяются для разработки устройств эффективного управления и обработки информации.

Список литературы

1. Чебурахин И.Ф. Синтез дискретных управляющих систем и математическое моделирование: алгоритмы, программы. - М.: Физматлит, 2004.
2. Чебурахин И.Ф.Математические модели для интеллектуализации синтеза дискретных логических управляющих устройств на основе цифровых интегральных схем // Изв. РАН. ТиСУ. 2008. № 1. - С. 68-77.
3. Чебурахин И.Ф. О минимизации сложности представления булевых функций из некоторых классов «Интеллек-

туальные системы и компьютерные науки» // X Междунар. конф.: Сб. материалов / Под общ. ред. В.А. Садовничего, В.Б. Кудрявцева, А.В. Михалева. - М.: МАКС Пресс, 2011. - С. 311-314.
4. Чебурахин И.Ф. Алгоритмизация представления булевых функций формулами и схемами минимальной сложности в базисе Жегалкина // Мехатроника, автоматизация, управление. 2012. № 12. - С. 7-14.
5. Теряев Е.Д., Петрин К.В., Филимонов А.Б., Филимонов Н.Б. Агентные технологии в автоматизированных информационно-управляющих системах. Ч.1., Ч.II // Мехатроника, автоматизация, управление. 2010:. № 7. - С. 11-27; № 10. - С. 25-34.
6. Чебурахин И.Ф., Цурков В.И.Синтез дискретных логических устройств обработки информации на основе теории агентов. // Мехатроника, автоматизация, управление. 2011. № 3. - с. 27-34.

УДК 531.36
Ф.Г. Гаращенко, д-р техн. наук, проф.
В.Т. Матвієнко, канд. фіз.-мат. наук, доц.
Київський національний університет імені Т Шевченка, м.Київ, Україна
fedir47@gmail.com, matvienko.vt@gmail,com

МОДЕЛЮВАННЯ ТА ОПТИМІЗАЦІЯ ДИНАМІКИ ЗАРЯДЖЕНИХ ПУЧКІВ ЯК ЗАДАЧА ПРАКТИЧНОЇ СТІЙКОСТІ

В даний момент прискорювачі заряджених частинок широко використовуються в наукових дослідженнях та різних галузях народного господарства. Проектування прискорювачів добре розроблено на основі фізичних принципів прискорення і фокусування. Але зараз постає питання проектування прискорюючих систем з оптимальними характеристиками, які б дозволяли при одному і тому ж рівні витрат отримати пучки з більшою енергією, з більшою густиною тощо.

Аналіз показує, що задачі практичної стійкості і деякі задачі оптимального керування потоками частинок тісно пов'язані [2, 3, 5, 6]. Це дозволило застосувати алгоритми побудови областей практичної стійкості для розв'язування задач оптимального керування пучками. Оптимізація систем в структурно-параметричному класі дозволило розробити ряд алгоритмів отримання квазіоптимальних розв'язків, які зарекомендували свою працездатність в реальних установках.

При практичних реалізаціях задачі керування пучком траєкторій дуже складні. Моделювання динаміки пучка на ЕОМ з паралельним визначенням полів із рівнянь Максвела є дуже важким процесом з обчислювальної точки зору [8, 9]. Представляючи поля в структурно-параметричному класі і, застосовуючи метод послідовного ускладнення математичної моделі, приходимо до оптиміза-

ційних задач, які можна чисельно розв'язати на ЕОМ. Цей підхід має такі позитивні аспекти:

а) аналіз фізики процесу і вдалий вибір параметрів оптимізації дозволяє отримати кращі локальні екстремуми функції цілі, що мінімізується, в заданому класі керуючих полів;

б) параметричне представлення полів дає можливість визначити оптимальні режими у структурах, які фізично можна реалізувати.

Задачі оптимального керування пучками розглядалися багатьма вченими. Зокрема таким проблемам присвячено праці [2-6, 11, 12].

Основою для розв'язування таких задач служать класичні методи, які містяться в монографіях [4 ,13].

Рівняння руху заряджених частинок в електромагнітних полях, їх аналіз

Приведемо рівняння руху заряджених частинок в електромагнітних полях. Швидкість зміни імпульсу частинки $p = m * \upsilon$ дорівнює силі Лоренца [9], що діє на неї :

$$\frac{dp}{dt} = Z \cdot e \cdot \{E + [\upsilon \cdot B]\}, \tag{1}$$

де Z – зарядове число, e – заряд частинки, E – вектор напруженості прискорюючого поля, B – вектор магнітної індукції, υ – швидкість частинок, $[\upsilon \cdot B]$ - векторний добуток,

$$m = \frac{m_0}{\sqrt{1 - \frac{\upsilon^2}{c^2}}} = \gamma \cdot m_0 \ , \tag{2}$$

$\gamma = \frac{m}{m_0}$ - приведена енергія частинки.

Запишемо (1) у вигляді:

$$m_0 \cdot \gamma \cdot \frac{d\upsilon}{dt} + m_0 \cdot \upsilon \cdot \frac{d\gamma}{dt} = Z \cdot e \cdot \{E + [\upsilon \cdot B]\} \cdot \tag{3}$$

Проектуючи (1) на координатні осі, отримаємо :

$$\begin{aligned}
\frac{d^2x}{dt^2} &= \frac{Ze}{m_0\gamma}\left[E_x + \frac{dy}{dt}B_z - \frac{dz}{dt}B_y - \frac{m_0}{Ze}\frac{dx}{dt}\frac{d\gamma}{dt}\right],\\
\frac{d^2y}{dt^2} &= \frac{Ze}{m_0\gamma}\left[E_y + \frac{dz}{dt}B_x - \frac{dx}{dt}B_z - \frac{m_0}{Ze}\frac{dy}{dt}\frac{d\gamma}{dt}\right],\\
\frac{d^2z}{dt^2} &= \frac{Ze}{m_0\gamma}\left[E_z + \frac{dx}{dt}B_y - \frac{dy}{dt}B_x - \frac{m_0}{Ze}\frac{dz}{dt}\frac{d\gamma}{dt}\right].
\end{aligned} \qquad (4)$$

З урахуванням сил кулонівської взаємодії (1) перепишемо у вигляді:

$$\begin{aligned}
\frac{d^2x}{dt^2} &= \frac{Ze}{m_0\gamma}\left[E_x + \frac{dy}{dt}B_z - \frac{dz}{dt}B_y - \frac{m_0}{Ze}\frac{dx}{dt}\frac{d\gamma}{dt} - \frac{1}{\gamma^2}\frac{\partial u}{\partial x}\right],\\
\frac{d^2y}{dt^2} &= \frac{Ze}{m_0\gamma}\left[E_y + \frac{dz}{dt}B_x - \frac{dx}{dt}B_z - \frac{m_0}{Ze}\frac{dy}{dt}\frac{d\gamma}{dt} - \frac{1}{\gamma^2}\frac{\partial u}{\partial y}\right],\\
\frac{d^2z}{dt^2} &= \frac{Ze}{m_0\gamma}\left[E_z + \frac{dx}{dt}B_y - \frac{dy}{dt}B_x - \frac{m_0}{Ze}\frac{dz}{dt}\frac{d\gamma}{dt} - \frac{1}{\gamma^2}\frac{\partial u}{\partial z}\right],
\end{aligned} \qquad (5)$$

де u – потенціал власного поля пучка, $\frac{d\gamma}{dt} = \frac{1}{\sqrt{1-\frac{\upsilon^2}{c^2}}}$.

Рівняння (5) є загальними рівняннями руху в часі. Виразимо похідну $\frac{d\gamma}{dt}$ через координати частинки і проекції її швидкості:

$$\frac{d\gamma}{dt} = \frac{d}{dt}\frac{1}{\sqrt{1-\frac{(\upsilon\cdot\upsilon)}{c^2}}} = \frac{\gamma^2}{c^2}\left(\upsilon\frac{d\upsilon}{dt}\right).$$

Визначаючи швидкість із (1) будемо мати :

$$\frac{d\gamma}{dt} = \frac{Ze}{m_0c^2}\left[E_x\frac{dx}{dt} + E_y\frac{dy}{dt} + E_z\frac{dz}{dt} - \frac{1}{\gamma^2}\left(\frac{dx}{dt}\frac{\partial u}{\partial x} + \frac{dy}{dt}\frac{\partial u}{\partial y} + \frac{dz}{dt}\frac{\partial u}{\partial z}\right)\right] . \qquad (6)$$

Припустимо, що зовнішнє поле змінюється по закону $E = E'\cos\varphi$, де E' - амплітуда напруженості прискорюючого поля, $\varphi = \frac{2\pi c}{\lambda}(t - t_B) - \varphi(0)$ - фаза прискорюючої хвилі, що діє на частинку в момент t, t_B - час поширення хвилі від початкової точки z=0 в точку z, $\varphi(0)$ - фаза прискорюючої хвилі, при якій іон вступає в процес прискорення, λ - довжина хвилі високочастотного прискорюючого поля. Отримані рівняння будемо розглядати відносно незалежної змінної z. Для цього запишемо:

$$(\upsilon,\upsilon) = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2 = \left(\frac{dz}{dt}\right)^2\left[1 + \left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2\right]$$

,

звідси

$$\left(\frac{dt}{dz}\right)^2 = \frac{1 + \left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2}{(\upsilon,\upsilon)}, \quad \frac{1}{\gamma^2} = 1 - \frac{(\upsilon,\upsilon)}{c^2},$$

$$(\upsilon,\upsilon) = \frac{(\gamma^2 - 1)c^2}{\gamma^2}.$$

Отже

$$\left(\frac{dt}{dz}\right)^2 = \frac{\gamma^2\left[1 + \left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2\right]}{(\gamma^2 - 1)c^2}, \qquad (7)$$

$$\frac{dx}{dt} = \frac{dx}{dz}\frac{dz}{dt}, \frac{d^2x}{dt^2} = \frac{d^2x}{dz^2}\left(\frac{d^2z}{dt^2}\right)^2 + \frac{dx}{dz}\frac{d^2z}{dt^2};$$

$$\frac{d^2x}{dt^2} = \left(\frac{d^2t}{dz^2}\right)^2\left[\frac{d^2x}{dz^2} + \frac{dx}{dz}\frac{d^2z}{dt^2}\right]. \qquad (8)$$

Використовуючи (8) з урахуванням (6) рівняння руху (5) можна записати у вигляді:

$$\frac{d^2x}{dz^2} =$$

$$= \frac{Ze\gamma}{m_0c^2\left(\gamma^2-1\right)}\left\{E_x - \frac{dx}{dz}E_z - \frac{1}{\gamma^2}\left(\frac{\partial u}{\partial x} - \frac{dx}{dz}\frac{\partial u}{\partial z}\right) + \left[\frac{dy}{dz}B_z - \left(1+\left(\frac{dx}{dz}\right)^2 B_y\right) + \right.\right.$$

$$\left.\left. + \frac{dx}{dz}\frac{dy}{dz}B_x\right]\frac{c\sqrt{\gamma^2-1}}{\gamma\sqrt{1+\left(\frac{dx}{dz}\right)^2+\left(\frac{dy}{dz}\right)^2}}\right\}\left[1+\left(\frac{dx}{dz}\right)^2+\left(\frac{dy}{dz}\right)^2\right],$$

$$\frac{d^2y}{dz^2} =$$

$$= \frac{Ze\gamma}{m_0c^2\left(\gamma^2-1\right)}\left\{E_y - \frac{dy}{dz}E_z - \frac{1}{\gamma^2}\left(\frac{\partial u}{\partial x} - \frac{dy}{dz}\frac{\partial u}{\partial z}\right) + \left[\left(1+\left(\frac{dy}{dz}\right)^2 B_x\right) - \frac{dx}{dz}B_z - \right.\right.$$

$$\left.\left. - \frac{dx}{dz}\frac{dy}{dz}B_y\right]\frac{c\sqrt{\gamma^2-1}}{\gamma\sqrt{1+\left(\frac{dx}{dz}\right)^2+\left(\frac{dy}{dz}\right)^2}}\right\}\left[1+\left(\frac{dx}{dz}\right)^2+\left(\frac{dy}{dz}\right)^2\right],$$

(9)

$$\frac{d\gamma}{dz} = \frac{Ze}{m_0c^2}\left[E_x\frac{dx}{dz} + E_y\frac{dy}{dz} + E_z - \frac{1}{\gamma^2}\left(\frac{dx}{dz}\frac{\partial u}{\partial x} + \frac{dy}{dz}\frac{\partial u}{\partial y} + \frac{\partial u}{\partial z}\right)\right]$$

,

$$\frac{d\varphi}{dz} = \frac{2\pi}{\lambda}\frac{\gamma}{\sqrt{\gamma^2-1}}\sqrt{1+\left(\frac{dx}{dz}\right)^2+\left(\frac{dy}{dz}\right)^2}\,.$$

Для (9) реалізація чисельного моделювання на ЕОМ процедур оптимізації є дуже складною. Тому розглянемо спрощену модель [8,11]. Рівняння руху (9) будемо розгля-

дати без врахування сил кулонівської взаємодії, тобто divE=0. В цьому випадку рівняння запишуться у вигляді:

$$\frac{d\gamma}{dz} = \frac{Ze}{m_0 c^2} E_z \cos\varphi,$$

$$\frac{d\varphi}{dz} = \frac{2\pi}{\lambda} \frac{\gamma}{\sqrt{\gamma^2 - 1}}, \qquad (10)$$

$$\frac{d^2x}{dz^2} = \frac{Ze}{m_0 c^2} \frac{\gamma}{\gamma^2 - 1} \left[\left(-\frac{1}{2} \frac{\partial E_z}{\partial z} + g(z) \right) x - E_z \frac{dx}{dz} \right] \cos\varphi,$$

$$\frac{d^2y}{dz^2} = \frac{Ze}{m_0 c^2} \frac{\gamma}{\gamma^2 - 1} \left[\left(-\frac{1}{2} \frac{\partial E_z}{\partial z} - g(z) \right) y - E_z \frac{dy}{dz} \right] \cos\varphi$$

,

где

$$g(z) = \frac{1}{2} \left[\frac{\partial E_x(0,0,z)}{\partial x} - \frac{\partial E_y(0,0,z)}{\partial y} \right].$$

Введемо безрозмірну координату $\xi = \frac{z}{\lambda}$ [11]. Тоді рівняння (10) приймуть вигляд:

$$\frac{d\gamma}{d\xi} = \alpha(\xi) \cos\varphi,$$

$$\frac{d\varphi}{d\xi} = \frac{2\pi\gamma}{\sqrt{\gamma^2 - 1}}$$

$$\frac{d^2x}{d\xi^2} = \frac{\gamma}{\gamma^2 - 1} \left[\left(-\frac{1}{2} \frac{d\alpha}{d\xi} + g(\xi) \right) x - \alpha(\xi) \frac{dx}{d\xi} \right] \cos\varphi, \qquad (11)$$

$$\frac{d^2y}{d\xi^2} = \frac{\gamma}{\gamma^2 - 1} \left[\left(-\frac{1}{2} \frac{d\alpha}{d\xi} - g(\xi) \right) y - \alpha(\xi) \frac{dy}{d\xi} \right] \cos\varphi.$$

Задачу вибору функцій керування $\alpha(\xi) = \frac{Ze\lambda Ez}{m_0 c^2}$, яка відповідає амплітуді напруженості прискорюючого поля при резонансному прискоренні та функції g(ξ) називають "E-g" задачею.

Постановки задач

Аналіз різних характеристик прискорювачів заряджених частинок показує, що потенціальні можливості установок для отримання оптимальних вихідних характеристик використовуються не повністю. Для отримання таких характеристик, наприклад, в прискорювачах з дрейфовими трубками важливо визначити положення трубок (їх довжину, форму), при яких досягається максимальна інтенсивність пучка з заданими розкидами його кінцевого стану.

Задача оптимізації пучка іонів в лінійному прискорювачі являє собою дуже складну проблему. Тому дану задачу при її дослідженні можна розбити на декілька напрямків : розробка методів розрахунку оптимальних прискорюючих структур; радіальне фокусування іонів; врахування власного електростатичного поля пучка; розробка швидкодіючих чисельних алгоритмів розрахунку зовнішніх електромагнітних полів і т.д.

Сформулюємо загальні постановки деяких задач, які в комплексі вирішують вихідну задачу. Кожна з цих задач має окремий інтерес і має самостійне значення[3,5,6,11,12].

Задача 1. Для поздовжнього руху при заданому початковому розкиді по енергії та фазі визначити структуру прискорювача таким чином, щоб на виході енергетичний та фазовий розкиди пучка були мінімальні.

Задача 2. Визначити параметри прискорюючої структури таким чином, щоб захват іонів в процес прискорення по фазі та енергії був максимальним при заданому енерге-

тично-фазовому розкиді в кінці прискорювача.

Задача 3. Для радіального руху з заданим фазовим об'ємом знайти параметри прискорювача таким чином, щоб на виході був мінімальний розкид по радіальним складовим та їх швидкостям.

Задача 4. При заданому розкиді в кінці прискорювача підібрати розміри трубок дрейфу і величини прискорюючих та фокусуючих полів таким чином, щоб захват частинок за початковими координатами та їх швидкостями був максимальним.

Задача 5. Сформульовані вище задачі дослідити з врахуванням того, що компоненти електричного поля задовольняють рівняння Лапласа. При цьому компоненти вектора напруженості прискорюючого поля виражаються через параметри, що оптимізуються.

Задача 6. Оптимізація динаміки пучка з врахуванням сил кулонівської взаємодії. В якості критерію якості розглянути мінімізацію розкиду координат пучка в кінці прискорювача, а також визначити максимальну область захвату іонів в процес прискорення.

Вибір початкових наближень в "E – g " задачі

Запишемо рівняння руху заряджених частинок в електромагнітних полях. Будемо розглядати несиметричний випадок, тобто

$$\begin{aligned}
&\frac{d\gamma}{d\xi} = \alpha(\xi)\cos\varphi,\\
&\frac{d\varphi}{d\xi} = \frac{2\pi\gamma}{\sqrt{\gamma^2-1}},\\
&\frac{d^2x}{d\xi^2} = \frac{\gamma}{\gamma^2-1}\left[\left(-\frac{1}{2}\frac{d\alpha}{d\xi}+g(\xi)\right)x-\alpha(\xi)\frac{dx}{d\xi}\right]\cos\varphi,\\
&\frac{d^2y}{d\xi^2} = \frac{\gamma}{\gamma^2-1}\left[\left(-\frac{1}{2}\frac{d\alpha}{d\xi}-g(\xi)\right)y-\alpha(\xi)\frac{dy}{d\xi}\right]\cos\varphi.
\end{aligned} \qquad (12)$$

Сформулюємо “E-g” задачу. Знайти вигляд функцій $g(\xi)$ та $\alpha(\xi)$ таким чином, щоб пучок в кінці прискорювача був сфокусованим за рахунок фокусуючого поля, тобто функції $g(\xi)$.

Згідно рівняння (10) функція $g(\xi)$ визначається таким чином:

$$g(z) = \frac{1}{2}\left[\frac{\partial E_x(0,0,z)}{\partial x} - \frac{\partial E_y(0,0,z)}{\partial y}\right].$$

Тобто, як видно з запису, функція $g(\xi)$ задає асиметрію прискорюючого поля. Похідні $\frac{\partial E_x(0,0,z)}{\partial x}$, $\frac{\partial E_y(0,0,z)}{\partial y}$ задають швидкість відхилення складових E_x та E_y поля відповідно від осей OX та OY, а їх різниця дає саме цю асиметрію: перевагу одного напрямку над іншим. Беручи до уваги вище сказане, зрозуміло, що впливаючи на цю різницю можна добитися фокусування.

Запишемо деякі критерії, що можна розглядати при розв’язуванні цієї задачі.

Задача 1. Мінімізація відхилення фазових координат.

Критерій для цієї постановки задачі має вигляд :

$$\Phi(x,y,g) = \min_{g}\left\{x^2(T) + y^2(T)\right\}. \qquad (13)$$

Задача 2. Максимізація інтервалу початкового значення фази.

Критерій має вигляд :

$$\Phi(\cdot) = \min_{\alpha,g}\ \max_{\varphi(0)\in[a,b]}\left\{x^2(T) + y^2(T)\right\}. \qquad (14)$$

В цьому випадку ми мінімізуємо відхилення фазових координат, але при цьому захват частинок в процес прискорення відбувається з якомога ширшим діапазоном

початкового значення фази.

Крім того на фазові координати можуть накладатися обмеження вигляду [3] :

$$x, y \in \Gamma_\xi = \left\{x, y : |x| \le \bar{x}(\xi), |y| \le \bar{y}(\xi)\right\}, \quad \xi \in [0, T]. \qquad (15)$$
$$x, y \in \Phi_\xi = \left\{x, y : x^2 + y^2 \le \bar{r}^2(\xi)\right\}$$

Список литературы

1. Бахвалов Н.С. Численные методы. – М.:Наука, 1973. – 631с.
2. Башняков О.М., Гаращенко Ф.Г., Пічкур В.В. Практична стійкість та структурна оптимізація динамічних систем. – К.: Видавничо– поліграфічний центр "Київський університет".– 197 с.
3. Бублик Б.Н., Гаращенко Ф.Г., Кириченко Н.Ф. Структурно – параметрическая оптимизация и устойчивость динамики пучков. – К.: Наукова думка, 1985. – 304 с.
4. Бублик Б.Н., Кириченко Н.Ф. Основы теории управления. – К.: Вища школа, 1975. –328 с.
5. Гаращенко Ф.Г. Применение методов практической устойчивости к моделированию и управлению пучками заряженных частиц // Автоматика. 1982. № 5. - С. 42-48.
6. Гаращенко Ф.Г. Иследование задач практической устойчивости численными методами и оптимизация динамики пучков // Прикладная математика и механіка. 1987. Т. 51, №5. - С. 717-723.
7. Калиткин Н.Н. Численные методы. – М.: Наука, 1978. – 512с.
8. Капчинский И.М. Динамика частиц в линейных резонансных ускорителях. – М.: Атомиздат, 1966. – 310 с.
9. Ландау Л.Д., Лифшиц Е.М. Теория поля. – М.: Наука, 1967. – 460 с.
10. Ляшко И.И., Макаров В.Л., Скоробогатько А.А. Методы вычислений. – К.: Вища школа, 1977. – 408 с.
11. Овсянников Д.А. Математические методы управления пучками. – Л.: Изд-во Ленингр.ун-та., 1980. – 228 с.

12. Овсянников Д.А. Моделиравание и оптимизация динамики пучков. – Л., 1990. – 310с.
13. Понтрягин Л.С., Болтянский В.Г, Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. – М.: Наука, 1976. – 392с.
14. Горбунов В.К. Метод параметризации задач оптимального управления // Журнал вычислительной математики и математической физики. 1979. Т. 19, №2. – С. 292-303.
15. Гаращенко Ф.Г. Недифференцируемые задачи структурно-параметрической оптимизации и проектирование ускоряющих и фокусирующих систем // Автоматика. 1986. №1. – С.50-53.
16. Гаращенко Ф.Г., Пичкур В.В. Структурная оптимизация динамических систем на основе обобщенного принципа Беллмана // Проблемы управления и информатики. 1997. № 6. – С. 6-13.
17. Гаращенко Ф.Г. Задачи структурной параметрической оптимизации разрывных динамических систем // Автоматика. 1985. № 2. – С. 37-42.
18. Емельянов С.В. Системы автоматического управления с переменной структурой. – М.: Наука, 1976. – 335 с.
19. Ладиков-Роев Ю.П., Самойленко Ю.И. Структурная оптимизация регулирующих сред // Проблемы управления и информатики. 1996. № 1-2. – С. 101-108.
20. Яковлев О.С. Метод структурного синтеза нелинейных регуляторов // Проблемы управления и информатики. 1996. № 1-2. – С. 211-223.
21. Sirisena M.K. A gradient method for computing optimal bang-bang controls // Int. J. of Contr. 1974. Vol. 19, № 2. – P. 257-264.

УДК 517.934
В.Т. Матвиенко, канд. физ.-мат. наук, доц.
Киевский национальный университет им. Тараса Шевченко, г. Киев, Украина
matvienko.vt@gmail.com

ОПТИМИЗАЦИЯ МОДАЛЬНЫХ РЕГУЛЯТОРОВ

Рассматривается задача оптимального выбора структуры распределения управляющего сигнала в линейной системе, с целью минимизации нормы матрицы коэффициентов усиления, в обратной связи закона модального регулирования.

Пусть в системе

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t), \tag{1}$$

где x - n - мерный, u - m мерный векторы, необходимо определить обратную связь

$$u(t) = Cx(t) \tag{2}$$

согласно условию модального управления

$$\det(\lambda E_n - A - BC) = \lambda^n + a_1\lambda^{n-1} + \ldots + a_n \tag{3}$$

при оптимизации $\|C\|$ за счет выбора элементов матрицы C.

Методы определения матрицы усиления модального регулятора приводятся в работах [1-5]. Один из возможных методов нахождения матрицы модального регулятора сводится к представлению искомой матрицы в виде $C = vq^T$ [4, 5]. Это представление матрицы усиления сужает множество возможных модальных регуляторов, но дает возможность сравнительно просто определить коэффициенты модального регулятора. В работе предложен следующий

подход по определению матрицы C. Представим систему (1) в виде

$$\frac{dx(t)}{dt} = Ax(t) + b_1u_1(t) + b_2u_2(t) + \ldots + b_mu_m(t) =$$

$$= (A + b_1c_1^T + b_2c_2^T + \ldots + b_mc_m^T)x(t),$$

где

$$C = \begin{pmatrix} c_1^T \\ \vdots \\ c_m^T \end{pmatrix}, B = (b_1, \ldots, b_m), u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}.$$

Вначале рассмотрим систему

$$\frac{dx(t)}{dt} = Ax(t) + b_1u_1(t)$$

и определим коэффициенты характеристического уравнения

$$\det(\lambda E_n - A - b_1c_1^T) = 0$$

по формуле [3]

$$p(1) = p(0) - P(0)S^T(b_1, A(0))c_1,$$

где

$$p(0) = p,\ A(0) = A,$$

элементами вектора p являются коэффициентами характеристического уравнения разомкнутой системы

$$\det(\lambda E_n - A) = \lambda^n + p_1\lambda^{n-1} + \ldots + p_n.$$

На следующем шаге рассматривается система

$$\frac{dx(t)}{dt} = (Ax(t) + b_1c_1^T)x(t) + b_2u_2(t)$$

с коэффициентами характеристического уравнения замк-

нутой системы

$$p(2) = p(1) - P(1)S^T(b_2, A(1))c_2,$$

где

$$A(1) = A(0) + b_1 c_1^T.$$

На шаге m рассматривается следующая система уравнений

$$\frac{dx(t)}{dt} = A(m-1)x(t) + b_m u_m(t),$$

где

$$A(m-1) = A(m-2) + b_{m-1} u_m(t),$$

для замкнутой системы, которой необходимо обеспечить следующие коэффициенты характеристического уравнения

$$p(m) = a = p(m-1) - P(m-1)S^T(b_m, A(m-1))c_m,$$

выбирая соответствующим образом вектор c_m [3].

Таким образом, в случае ограничений вида $c_j \in \Omega_C(j),$ рассматриваемая задача синтеза с оптимизацией модального регулятора сводится к следующей задаче управления системой с дискретным аргументом

$$\begin{aligned} p(k+1) &= p(k) - P(k)S^T(b_{k+1}, A(k))c_{k+1}, \\ A(k+1) &= A(k) + b_{k+1}c_{k+1}^T,\ k = 0,1,2,\ldots,m-1 \end{aligned} \tag{4}$$

из начального состояния

$$p(0) = p,\ A(0) = A \tag{5}$$

в финальное

$$p(m) = a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{6}$$

при условии оптимизации

$$I(C) = \min_{c_j \in \Omega_{\mathbf{c}}(j), \forall j = \overline{1,m}} \sum_{j=1}^{m} \left\| c_j \right\|^2 + \left\| p(m) - a \right\|. \tag{7}$$

Здесь вектор $p^T = (p_1, \ldots, p_n)$ определяется условием

$$\det(\lambda E_n - A) = \lambda^n + p_1 \lambda^{n-1} + \ldots + p_n,$$

$$P(k) = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ p_1(k) & 1 & 0 & \ldots & 0 \\ p_2(k) & p_1(k) & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ p_{n-1}(k) & p_{n-2}(k) & p_{n-3}(k) & \ldots & 1 \end{pmatrix},$$

$$S(b_{k+1}, A(k)) = (b_{k+1}, \ldots, A^{n-1}(k) b_{k+1}).$$

Для решения поставленной задачи запишем численную процедуру нахождения матрицы C. С этой целью выпишем функцию Гамильтона [6] для системы (4)

$$H(p(k), A(k), \Psi(k+1), \Phi(k+1), k) = - \left\| c_{k+1} \right\|^2 +$$

$$+ \Psi^T(k+1) \left[p(k) - P(k) S^T(b_{k+1}, A(k)) c_{k+1} \right] + tr \left[\Phi^T(k+1) (A(k) + b_{k+1} c_{k+1}^T) \right].$$

Сопряженные переменные $\Psi(k), \Phi(k)$ удовлетворяют следующим системам уравнений

$$\Psi(k) = grad_{p(k)} H(\cdot) =$$

$$= \left[E_n - \begin{pmatrix} 0 & c_{k+1}^T b_{k+1} & c_{k+1}^T A(k) b_{k+1} & \dots & c_{k+1}^T A^{n-2}(k) b_{k+1} \\ 0 & 0 & c_{k+1}^T b_{k+1} & \dots & c_{k+1}^T A^{n-3}(k) b_{k+1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \cdots & c_{k+1}^T b_{k+1} \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \right] \Psi(k+1)$$

,

$$\Psi(m) = 0,\ k = m-1,\dots,0,$$

$$\Phi(k) = grad_{A(k)} H(\cdot) =$$

$$= \Phi(k+1) - \left(c_{k+1}, A^T(k) c_{k+1}, \dots, A^{T^{n-2}}(k) c_{k+1}\right) \times$$

$$\times \left(\overline{E}_2 P^T(k) \Psi(k+1), \overline{E}_3 P^T(k) \Psi(k+1), \dots, \overline{E}_n P^T(k) \Psi(k+1)\right) \begin{pmatrix} b_{k+1}^T \\ b_{k+1}^T A^T(k) \\ \vdots \\ b_{k+1}^T A^{T^{n-2}}(k) \end{pmatrix},$$

$$\Phi(m) = 0,\ k = m-1,\dots,0.$$

Матрицы $\overline{E}_i, i = \overline{2,n}$ размерности $(n-1) \times n$ имеют следующую структуру

$$\overline{E}_2 = \begin{pmatrix} e_2^T \\ e_3^T \\ \vdots \\ e_n^T \end{pmatrix}, \overline{E}_3 = \begin{pmatrix} e_3^T \\ e_4^T \\ \vdots \\ 0 \end{pmatrix}, \dots, \overline{E}_n = \begin{pmatrix} e_n^T \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

где $e_i - i = \overline{2,n}$ единичные орты размерности *n*. Тогда

$$grad_{\mathbf{c}_{i+1}} I(C) = - grad_{\mathbf{c}_{i+1}} H(\cdot) =$$
$$= 2c_{i+1} + S(b_{i+1}, A(i)) P^T(i) \Psi(i+1) - \Phi^T(i+1) b_{i+1},$$

для градиентных вычислительных процедур

$$\overline{c_j} = c_j - \rho_j grad_{\mathbf{c}_j} I(C),$$
$$j = 1,2,\ldots,m.$$

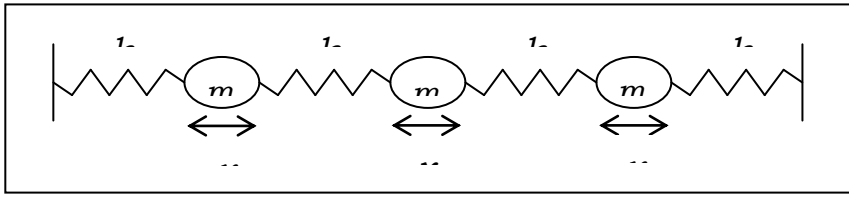


Рисунок 1

Был проведен численный эксперимент для колебательной системы с тремя массами (см. рисунок 1), которые закреплены пружинами между собой и управляемые двумя управляющими воздействиями.

Колебания центр масс такой системы описывается следующей системой

$$\frac{dx}{dt} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{k_1 + k_2}{m_1} & 0 & \frac{k_2}{m_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{k_2}{m_2} & 0 & -\frac{k_1 + k_2}{m_2} & 0 & \frac{k_3}{m_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{k_3}{m_3} & 0 & -\frac{k_3 + k_4}{m_3} & 0 \end{pmatrix} x + \begin{pmatrix} 0 & 0 \\ \frac{1}{m_1} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{m_1} \end{pmatrix} u, \quad (8)$$

где вектор состояния системы колебания имеет вид $x^T = (x_1, x_2, \ldots, x_6)$, x_1, x_2, x_3 – координаты положе-

ний центров масс m_1, m_2, m_3, соответственно, k_1, k_2, k_3 – коэффициенты жесткости соответствующих пружин. Управляющие силы для системы (8) приложены к массам m_1 и m_3

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}.$$

Замкнутой системе (8) необходимо обеспечить следующие собственные значения

$$\lambda_1 = -0.25 + 0.1i, \ \lambda_2 = -0.25 - 0.1i,$$

$$\lambda_3 = -0.2 + 0.1i, \ \lambda_4 = -0.2 - 0.1i,$$

$$\lambda_5 = -0.1 + 0.1i, \ \lambda_6 = -0.1 - 0.1i,$$

которые обеспечат замкнутой системе асимптотическую устойчивость с минимальной по норме модальной матрицей. Массы и коэффициенты жесткости были выбраны следующие:

$$m_1 = 12, \ m_2 = 25, \ m_3 = 8,$$
$$k_1 = 0.14, \ k_2 = 0.9, \ k_3 = 0.7, \ k_4 = 0.12.$$

В результате проведения числового эксперимента для колебательной системы (8), получена следующая оптимальная матрица модального управления

$$C_{опт} = \begin{pmatrix} -0.5467 & -1.0999 & -5.7125 & -1.7209 & 0.1374 & -0.3789 \\ 0.4998 & 0.0114 & 1.2698 & -0.4802 & 0.2817 & -0.0001 \end{pmatrix}$$

.

На рисунке 2 изображено колебание центров масс системы без управляющих воздействий.

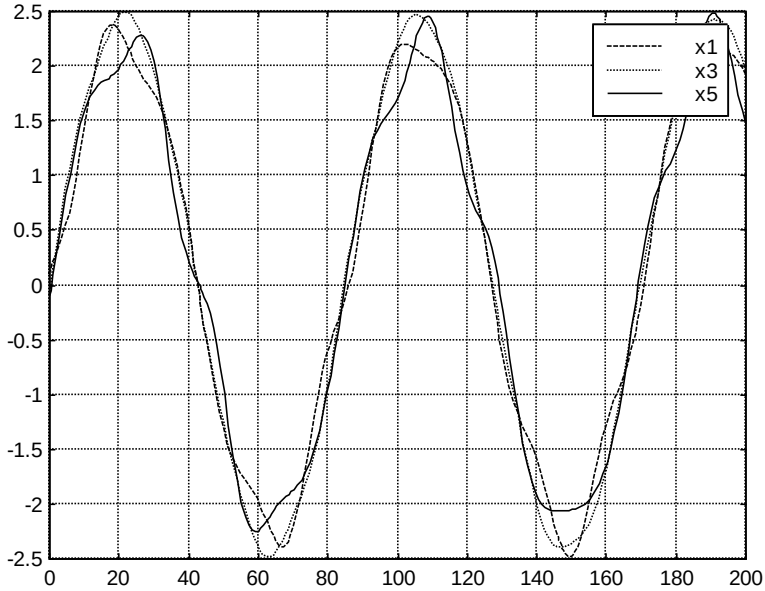


Рисунок 2

На рисунке 3 изображено колебание центров масс замкнутой системы с оптимальным модальным регулятором.

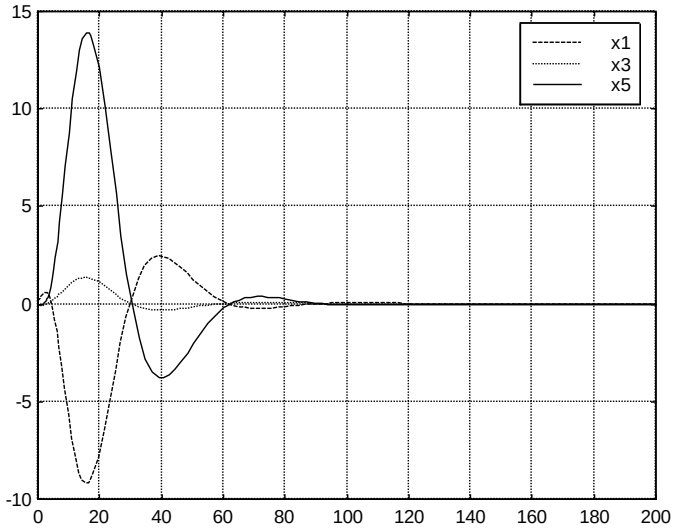


Рисунок 3

На рисунке 4 отображен процесс сходимости нормы матрицы модального регулятора к его оптимальному значению.

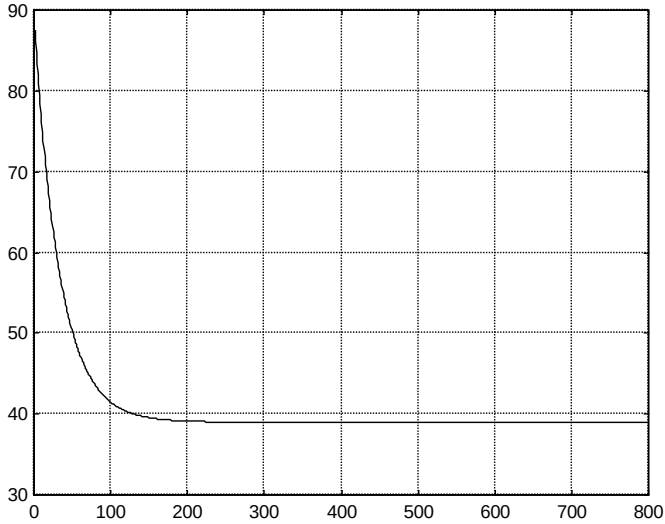


Рисунок 4

Список литературы

1. Андреев Ю.Н. Управление конечномерными линейными объектами. – М.: Наука, 1976.
2. Кузовков Н.Т. Модальное управление и наблюдающие устройства. – М.: Машиностроение, 1976.
3. Кириченко Н.Ф. Введение в теорию стабилизации движения. – К.: Вища школа, 1978.
4. Кириченко Н.Ф., Матвиенко В.Т. Оптимизация в задаче модального управления // В кн. "Вопросы оптимизации в динамических системах с непрырывно-дискретными параметрами". – К.: Наукова думка, 1980. – С .27–34.
5. Кириченко Н.Ф., Матвиенко В.Т. Оптимальный синтез структур для линейных систем управления // Проблемы управления и информатики. 1996. № 1-2. – С. 162-171.
6. Пропой Ф.И. Элементы теории оптимальных дискретных процессов. – М.: Наука, 1973.

УДК 681.5.09
Ю.Е. Обжерин, д-р техн. наук, проф.
Севастопольский национальный технический университет, г. Севастополь, Украина

ПОЛУМАРКОВСКИЕ МОДЕЛИ СИСТЕМ С РЕЗЕРВОМ ВРЕМЕНИ

Временное резервирование представляет собой название метода обеспечения нормального функционирования систем, выполняющих определенные задачи в условиях воздействий внешних возмущений, путем назначения и использования резервного (избыточного) времени [1,2]. В отличие от других видов избыточности здесь резервом является время. Отказ системы с резервом времени возникает в момент израсходования резерва времени.

Временное резервирование тесно связано с использованием при решении задач планирования и управления производством следующего принципа локализации [3]: если по истечении некоторого периода времени изменения интенсивностей затрат - выпуска производственного звена оказываются сбалансированными, а в течении этого периода запасы на входе и выходе звена не выходят за ограничения, то эти изменения не оказывают никакого влияния на условия работы остальной части системы и условия работы данного звена за пределами периода.

В этом принципе проявляется роль звеньев хранения как "буферных" элементов, ограничивающих распространение возмущений.

Источники резерва времени могут быть различными. В производственном процессе резерв времени может быть создан за счет внутренних запасов выходной продукции. Это характерно для многофазных автоматизированных производств с промежуточными накопителями. При разработке таких систем возникает ряд важных задач, связанных с выбором способа компоновки элементов системы

до и после накопителя, запаса производительности, емкости накопителей.

Методы временного резервирования могут быть классифицированы по ряду признаков [1,2], в том числе по характеру пополнения и использования резерва времени.

По способу пополнения резерва времени различают пополняемый, непополняемый и комбинированный резерв времени. В первом случае после окончания восстановления работоспособности объекта резерв времени пополняется: увеличивается по определенному закону. Резерв времени также делится на мгновенно пополняемый и пополняемый постепенно. Значение непополняемого (кумулятивного) резерва времени устанавливается заранее, и в процессе работы он может лишь уменьшаться за счет потерь времени на восстановление работоспособности объекта и устранение последствий отказов. Комбинированный резерв времени предполагает совместное использование пополняемого и непополняемого резерва времени. В этом случае порядок использования составляющих резерва времени зависит от алгоритмов функционирования объекта.

В работах [5-8] построены полумарковские модели с дискретно-непрерывным фазовым пространством состояний ряда систем с резервом времени.

В данном докладе рассматривается полумарковская модель двуфазной системы с промежуточным накопителем, структура которой изображена на рисунке 1.

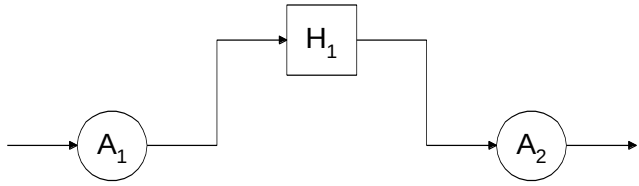


Рисунок 1 - Структурная схема двухфазной системы с промежуточным накопителем

Опишем порядок функционирования системы. На вход ее первого устройства A_1 производительностью c_1 на каждом отрезке времени Δt поступает столько единиц продукции $N_1(\Delta t)$, сколько оно может обработать за это время при безотказной работе, т.е. $N_1(\Delta t) = c_1 \Delta t$. После обработки на первом устройстве продукция поступает на второе устройство A_2, $N_2(\Delta t) = c_2 \Delta t$. Время безотказной работы (восстановления) устройства $A_i, i = \overline{1,2}$, является СВ $\alpha_i^{(0)}(\alpha_i^{(1)})$ с ФР $F_i^{(0)}(x)(F_i^{(1)}(x))$. Предполагается, что СВ $\alpha_i^{(0)}, \alpha_i^{(1)}$, независимы, имеют конечные математические ожидания; у ФР $F_i^{(0)}(x), F_i^{(1)}(x)$ существуют плотности $f_i^{(0)}(x), f_i^{(1)}(x)$. Емкость накопителя H_1, который предполагается абсолютно надежным, выражается в единицах времени, которое понадобится устройству A_2 для полного освобождения накопителя; максимальная емкость накопителя равна $h \geq 0$. При отказе устройства A_2 и заполненном накопителе устройство A_1 отключается, а при отказе устройства A_1 и пустом накопителе отключается устройство A_2. В случае различной производительности и работоспособном состоянии обоих устройств быстродействующее устройство может работать, не используя полностью свою производительность, если накопитель заполнен полностью при $c_1 > c_2$ и, если пуст, при $c_1 < c_2$. Считается, что количество продукта является непрерывной величиной (дискретность, составляющих продукт изделий, необходимо учитывать при небольшой емкости накопителя). Восстановление устройств A_1, A_2 предполагается неограниченным. Вся система считается работоспособной в данный момент времени, если она выдает продукцию на вы-

ходе устройства A_2.

Таким образом, на выходе устройства A_1 имеет место раздельный резерв времени, который может пополняться либо за счет запаса по быстродействию устройства A_1, либо за счет простоя устройства A_2. Описанную систему будем условно называть реальной системой.

Предположим, что устройства A_1 и A_2 имеют не возрастающую производительность ($c_1 \geq c_2$).

Для описания функционирования реальной системы введем следующее пространство полумарковских состояний:

$$\mathrm{E} = \left\{ i\bar{d}\bar{x}z : i = \overline{1,2}, \bar{d} = (d_1, d_2), \bar{x} = (x_1, x_2) \right\},$$

где i – номер устройства A_i, отказавшего или восстановившегося последним; элемент d_k вектора $\bar{d}$ фиксирует состояние устройства A_k: работоспособное ($d_k = 0$), восстановления ($d_k = 1$), отключения ($d_k = 2$). Значением элемента x_k вектора $\bar{x}$ является время, прошедшее с момента последнего восстановления или отказа устройства $A_k, x_i = 0$. Значение z определяет величину резерва времени в накопителе $H_1, 0 \leq z \leq h$.

Для приближенного вычисления стационарных характеристик системы используем метод, предложенный в [4].

Можно полностью построить полумарковский процесс, описывающий функционирование реальной системы. В данном случае будут определены только те характеристики полумарковского процесса, которые используются в применяемом методе.

Выберем опорную систему S_0. Предположим, что у устройства A_1 быстрое восстановление, т.е. его время восстановления $\alpha_1^{(1)}$ зависит от малого положительного параметра ε таким образом, что

$$\lim_{\varepsilon \to 0} M\alpha_1^{(1,\varepsilon)} = 0, \qquad (1)$$

а у выходного устройства A_2 времена безотказной работы и восстановления фиксированы. Это приводит к тому, что опорной системой S_0 будет являться система, у которой устройство A_1 восстанавливается мгновенно, а накопитель H_1 полностью заполнен. Временная диаграмма функционирования опорной системы изображена на рисунке 2.

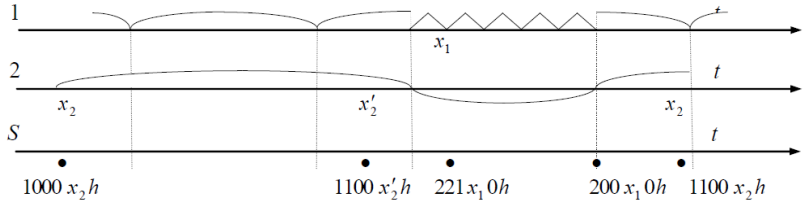


Рисунок 2 - Временная диаграмма функционирования опорной системы

Определим вероятности переходов вложенной цепи Маркова опорной системы:

$$p_{1000x_2h}^{1100y_2h} = \frac{f_1^{(0)}(y_2 - x_2)\overline{F}_2^{(0)}(y_2)}{\overline{F}_2^{(0)}(x_2)}, y_2 > x_2, \; p_{1000x_2h}^{221y_1 0h} = \frac{f_2^{(0)}(y_1 + x_2)\overline{F}_1^{(0)}(y_1)}{\overline{F}_2^{(0)}(x_2)}, y_1 > 0;$$

$$p_{200x_1 0h}^{221y_1 0h} = \frac{f_2^{(0)}(y_1 - x_1)\overline{F}_1^{(0)}(y_1)}{\overline{F}_1^{(0)}(x_1)}, y_1 > x_1, \; p_{200x_1 0h}^{1100y_2h} = \frac{f_1^{(0)}(y_2 + x_1)\overline{F}_2^{(0)}(y_2)}{\overline{F}_1^{(0)}(x_1)}, y_1 > 0;$$

$$\mathrm{P}_{1100x_2h}^{1000x_2h} = \mathrm{P}_{221x_1 0h}^{200x_1 0h} = 1.$$

Также как в работе [4] можно показать, что плотность стационарного распределения вложенной цепи Маркова опорной системы имеет следующий вид:

$$\begin{aligned}\rho(1000x_2h) = \rho(1100x_2h) = \rho_0\overline{F}_2^{(0)}(x_2),\\ \rho(200x_10h) = \rho(221x_10h) = \rho_0\overline{F}_1^{(0)}(x_1),\end{aligned} \quad (2)$$

где константа ρ_0 находится в явном виде из условия нормировки.

Таким образом, в класс эргодических состояний E^0 опорной системы S_0 входят состояния:

$$\mathrm{E}^0 = \{1000x_2h, 1100x_2h, 200x_10h, 221x_10h\}.$$

Для реальной системы с указанным выше критерием отказа системы подмножества работоспособных E_+ и отказовых состояний E_- имеют вид:

$$\begin{aligned}\mathrm{E}_+ &= \{1000\,x_2z, 1100\,x_2z, 200\,x_10z, 210\,x_10z\},\\ \mathrm{E}_- &= \{1010x_2z, 1110x_2z, 1120x_20, 121x_1x_2h,\\ &\quad 201x_10z, 211x_10z, 212x_1x_20, 221x_10h\}.\end{aligned}$$

Отметим, что состояния с дискретными кодами вида 102, 120, 122, 202, 220, 222 в описании функционирования опорной системы не используются.

Для приближенного нахождения средней стационарной наработке на отказ $\mathrm{T}_+^{(h)}$ и среднего стационарного времени восстановления $\mathrm{T}_-^{(h)}$ реальной системы в работе [4] предлагается использовать следующие формулы:

$$\mathrm{T}_+ \approx \frac{\int\limits_{E_+} m(x)\rho(dx)}{\int\limits_{E_+} P(x,E_-)\rho(dx)}, \quad \mathrm{T}_- \approx \frac{\int\limits_{E}\rho(dx)\int\limits_{E_-} m(y)P(x,dy)}{\int\limits_{E_+} P(x,E_-)\rho(dx)}. \quad (3)$$

Найдем выражения, входящие в формулы (3).

$$\int_{\mathrm{E}_+} \mathrm{P}(x,\mathrm{E}_-)\rho(dx) = \int_0^\infty \rho(1000x_2h)\mathrm{P}(1000x_2h,\mathrm{E}_-)dx_2 +$$

$$+\int_0^\infty \rho(1100x_2h)\mathrm{P}(1100x_2h,\mathrm{E}_-)dx_2 + \int_0^\infty \rho(200x_10h)\mathrm{P}(200x_10h,\mathrm{E}_-)dx_1 =$$

$$= \rho_0\left(\int_0^\infty \overline{F}_1^{(0)}(y_1)\overline{F}_2^{(0)}(y_1)dy_1 + \overline{F}_1^{(1,\varepsilon)}(h)\int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 +\right.$$

$$\left.+\int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)dy_1\int_0^\infty f_2^{(0)}(x_2+y_1)dx_2 + \int_0^\infty \overline{F}_1^{(0)}(y_1)dy_1\int_0^{y_1} f_2^{(0)}(y_1-x_1)dx_1\right) =$$

$$= \rho_0\left(M\alpha_1^{(0)} + \overline{F}_1^{(1,\varepsilon)}(h)\int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 + \int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)\overline{F}_2^{(0)}(y_1)dy_1\right).$$

Определим $\int_{\mathrm{E}_+} m(x)\rho(dx)$, для этого предварительно найдем средние времена пребывания в состояниях $m(x)$ реальной системы в эргодических состояниях:

$$\theta_{1000x_2h} = \alpha_1^{(0)} \wedge [\alpha_2^{(0)} - x_2]^+, m(1000x_2h) = \int_0^\infty \frac{\overline{F}_1^{(0)}(t)\overline{F}_2^{(0)}(x_2+t)}{\overline{F}_2^{(0)}(x_2)}dt,$$

$$\theta_{1100x_2h} = \alpha_1^{(1,\varepsilon)} \wedge [\alpha_2^{(0)} - x_2]^+ \wedge h, m(1100x_2h) = \int_0^h \frac{\overline{F}_1^{(1,\varepsilon)}(t)\overline{F}_2^{(0)}(x_2+t)}{\overline{F}_2^{(0)}(x_2)}dt,$$

$$\theta_{200x_1 0h} = \alpha_2^{(0)} \wedge [\alpha_1^{(0)} - x_1]^+, m(200x_1 0h) = \int_0^\infty \frac{\overline{F}_2^{(0)}(t)\overline{F}_1^{(0)}(x_1+t)}{\overline{F}_1^{(0)}(x_1)}\,dt.$$

Следовательно,

$$\int_{E_+} m(x)\rho(dx) = \rho_0\left(\int_0^\infty dx_2 \int_0^\infty \overline{F}_1^{(0)}(t)\overline{F}_2^{(0)}(x_2+t)dt + \right.$$

$$\left. + \int_0^\infty dx_2 \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)\overline{F}_2^{(0)}(x_2+t)dt + \int_0^\infty dx_2 \int_0^\infty \overline{F}_2^{(0)}(t)\overline{F}_1^{(0)}(x_1+t)dt \right) =$$

$$= \rho_0\left(M\alpha_1^{(0)} M\alpha_2^{(0)} + \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt \int_0^\infty \overline{F}_2^{(0)}(x_2+t)dx_2 \right).$$

Таким образом, средняя стационарная наработка на отказ $T_+^{(h)}$ рассматриваемой реальной системы приближенно вычисляется по формуле:

$$T_+^{(h)} \approx \left(M\alpha_1^{(0)} M\alpha_2^{(0)} + \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt \int_0^\infty \overline{F}_2^{(0)}(x_2+t)dx_2 \right) /$$

$$/ \left(M\alpha_1^{(0)} + \overline{F}_1^{(1,\varepsilon)}(h) \int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 + \int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)\overline{F}_2^{(0)}(y_1)dy_1 \right) . \qquad (4)$$

В силу условия (1)

$$\int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)\overline{F}_2^{(0)}(y_1)dy_1 = M(\alpha_1^{(1,\varepsilon)} \wedge \alpha_2 \wedge h) \approx M(\alpha_1^{(1,\varepsilon)} \wedge h),$$

$$\int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt \int_t^\infty \overline{F}_2^{(0)}(x_2)dx_2 = \int_0^h \overline{F}_2^{(0)}(x_2)dx_2 \int_0^{x_2} \overline{F}_1^{(1,\varepsilon)}(t)dt +$$

$$+ \int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 \int_0^h \overline{F}_2^{(1,\varepsilon)}(t)dt \approx$$

$$\int_0^h \overline{F}_2^{(0)}(x_2)dx_2 \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt +$$

$$+ \int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 \int_0^h \overline{F}_2^{(1,\varepsilon)}(t)dt = M\alpha_2^{(0)} \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt.$$

Поэтому, наряду с формулой (4), можно использовать следующее выражение для $\mathrm{T}_+^{(h)}$:

$$\mathrm{T}_+^{(h)} \approx \left(M\alpha_1^{(0)} M\alpha_2^{(0)} + M\alpha_2^{(0)} \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt \right) / \left(M\alpha_1^{(0)} + \right.$$

$$\left. + \overline{F}_1^{(1,\varepsilon)}(h) \int_h^\infty \overline{F}_2^{(0)}(t)dt + \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt \right), \tag{5}$$

а также

$$\mathrm{T}_+^{(h)} \approx (M\alpha_1^{(0)} M\alpha_2^{(0)}) \Big/ \left(M\alpha_1^{(0)} + \overline{F}_1^{(1,\varepsilon)}(h) \int_h^\infty \overline{F}_2^{(0)}(t)dt \right). \tag{6}$$

Перейдем к нахождению среднего стационарного времени восстановления $\mathrm{T}_-^{(h)}$. Знаменатель формулы (3) определен при вычислении $\mathrm{T}_+^{(h)}$, а числитель находится

следующим образом:

$$\int_{E}\rho(dx)\int_{E_-}m(y)\mathrm{P}(x,dy)=\rho_0\Bigg(M\alpha_2^{(1)}\int_0^\infty dx_2\int_0^\infty f_2^{(0)}(x_2+y_1)\overline{F}_1^{(0)}(y_1)dy_1+$$

$$+\int_0^\infty \overline{F}_2^{(0)}(x_2+h)dx_2\int_0^\infty \overline{F}_1^{(1,\varepsilon)}(h+t)dt+$$

$$+\int_0^\infty dx_2\int_0^h f_2^{(0)}(x_2+y_1)dy_1\int_0^\infty \overline{F}_2^{(1)}(t)\overline{F}_1^{(1,\varepsilon)}(y_1+t)dt+$$

$$+M\alpha_2^{(1)}\int_0^\infty dx_1\int_{x_1}^\infty f_2^{(0)}(y_1-x_1)\overline{F}_1^{(0)}(y_1)dy_1\Bigg)=\rho_0\Big(M\alpha_1^{(0)}M\alpha_2^{(1)}+$$

$$+\int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2\int_h^\infty \overline{F}_1^{(1,\varepsilon)}(t)dt+$$

$$\int_0^h \overline{F}_2^{(0)}(y_1)dy_1\int_0^\infty \overline{F}_2^{(1)}(t)\overline{F}_1^{(1,\varepsilon)}(y_1+t)dt\Big).$$

Следовательно,

$$\mathrm{T}_-^{(h)}\approx\left(M\alpha_1^{(0)}M\alpha_2^{(1)}+\int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2\int_h^\infty \overline{F}_1^{(1,\varepsilon)}(t)dt+M\alpha_2^{(1)}\int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)dy_1\right)/$$

$$/\left(M\alpha_1^{(0)}+\overline{F}_1^{(1,\varepsilon)}(h)\int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2+\int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)dy_1\right). \quad (7)$$

Учитывая, что

$$\int_0^h \overline{F}_2^{(0)}(y_1)dy_1 \int_0^\infty \overline{F}_2^{(1)}(t)\overline{F}_1^{(1,\varepsilon)}(y_1+t)dt = \int_0^\infty \overline{F}_2^{(1)}(t)dt \int_0^h \overline{F}_2^{(0)}(y_1)\overline{F}_1^{(1,\varepsilon)}(y_1+t)dy_1 \approx$$

$$\approx \int_0^\infty \overline{F}_2^{(1)}(t)dt \int_0^h \overline{F}_2^{(0)}(y_1)\overline{F}_1^{(1,\varepsilon)}(y_1)dy_1 \approx M\alpha_2^{(1)} \int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)dy_1,$$

получаем также

$$\mathrm{T}_-^{(h)} \approx \left(M\alpha_1^{(0)} M\alpha_2^{(1)} + \int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 \int_h^\infty \overline{F}_1^{(1,\varepsilon)}(t)dt + M\alpha_2^{(1)} \int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)dy_1 \right) /$$

$$/ \left(M\alpha_1^{(0)} + \overline{F}_1^{(1,\varepsilon)}(h) \int_h^\infty \overline{F}_2^{(0)}(x_2)dx_2 + \int_0^h \overline{F}_1^{(1,\varepsilon)}(y_1)dy_1 \right) \quad (8)$$

и

$$\mathrm{T}_-^{(h)} \approx \left(M\alpha_1^{(0)} M\alpha_2^{(1)} + \int_h^\infty \overline{F}_2^{(0)}(x)dx \int_h^\infty \overline{F}_1^{(1,\varepsilon)}(t)dt \right) /$$

$$/ \left(M\alpha_1^{(0)} + \overline{F}_1^{(1,\varepsilon)}(h) \int_h^\infty \overline{F}_2^{(0)}(t)dt \right). \quad (9)$$

Применяя формулу $K_\Gamma = \dfrac{T_+}{T_+ + T_-}$, определим стационарный коэффициент готовности реальной системы:

$$\mathrm{K}_\Gamma^{(h)} \approx \left(M\alpha_1^{(0)} M\alpha_2^{(0)} + M\alpha_2^{(0)} \int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt \right] / \left[M\alpha_1^{(0)} M\alpha_2^{(0)} + \right.$$

$$+M\alpha_1^{(0)}M\alpha_2^{(1)}+M\alpha_2^{(0)}\int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt+M\alpha_2^{(1)}\int_0^h \overline{F}_1^{(1,\varepsilon)}(t)dt+ \tag{10}$$

$$+\int_h^\infty \overline{F}_2^{(0)}(x)dx\int_h^\infty \overline{F}_1^{(1,\varepsilon)}(t)dt\Bigg),$$

а также

$$\mathrm{K}_\Gamma^{(h)} \approx (M\alpha_1^{(0)}M\alpha_2^{(0)})\Big/\Big(M\alpha_1^{(0)}M\alpha_2^{(0)}+M\alpha_1^{(0)}M\alpha_2^{(1)}+\int_h^\infty \overline{F}_2^{(0)}(x)dx\int_h^\infty \overline{F}_1^{(1,\varepsilon)}(t)dt\Bigg). \tag{11}$$

Используя стационарный коэффициент готовности $\mathrm{K}_\Gamma^{(h)}$, можно приближенно определить производительность $\Pi^{(h)}$ двухфазной системы с накопителем:

$$\Pi^{(h)} = \mathrm{K}_\Gamma^{(h)} \cdot c_2,$$

где c_2 – производительность выпускного устройства A_2.

Приведем пример использования формул (4) – (11). Пусть времена безотказной работы устройств A_1, A_2, одинаковой производительности имеют распределение Эрланга четвертого порядка с параметром *λ=0,1*, т.е. среднее время безотказной работы устройств равно 40 ч. Времена восстановления устройств имеют распределение Релея с параметром *σ=1,356,* среднее время восстановления устройств равно 1,7 ч.

В таблице 1 приведены значения $\mathrm{T}_+(h), \mathrm{T}_-(h), \mathrm{K}_\Gamma(h)$, вычисленные по формулам (4) – (11).

Таблица 1 - Характеристики надежности двухфазной системы с промежуточным накопителем

h	$T_{+}^{(h)}$ (5)	$T_{+}^{(h)}$ (6)	$T_{-}^{(h)}$ (8)	$T_{-}^{(h)}$ (9)	$K_{\Gamma}^{(h)}$ (10)	$K_{\Gamma}^{(h)}$ (11.)
0	20	20	1,7	1,7	0,922	0,922
0,1	20,077	20,052	1,652	1,625	0,924	0,924
0,2	20,209	20,159	1,61	1,609	0,926	0,926
0,3	20,394	20,32	1,571	1,571	0,928	0,928
0,4	20,633	20,535	1,537	1,536	0,931	0,93
0,5	20,926	20,805	1,507	1,506	0,933	0,932
0,6	21,272	21,128	1,482	1,48	0,935	0,935
0,7	21,67	21,505	1,46	1,458	0,937	0,937
0,8	22,121	21,935	1,442	1,44	0,939	0,938
0,9	22,621	22,417	1,428	1,425	0,941	0,94
1,0	23,17	22,949	1,417	1,417	0,942	0,942
1,1	23,766	23,53	1,41	1,406	0,944	0,944
1,2	24,406	24,157	1,406	1,401	0,946	0,945
1,3	25,087	24,826	1,405	1,4	0,947	0,947
1,4	25,804	25,535	1,407	1,401	0,948	0,948
1,5	26,552	26,278	1,412	1,406	0,95	0,949
1,6	27,327	27,05	1,419	1,412	0,951	0,95
1,7	28,122	27,844	1,428	1,421	0,952	0,951
1,8	28,93	28,655	1,439	1,432	0,953	0,952
1,9	29,744	29,475	1,452	1,445	0,953	0,953
2,0	30,556	30,295	1,466	1,459	0,954	0,954

Список литературы

1. Черкесов Г.Н. Надежность технических систем с временной избыточностью. - М.: Сов. радио, 1974. - 296 с.

2. Креденцер Б.П. Прогнозирование надежности систем с временной избыточностью. – К.: Наук. думка, 1978. - 240 с.

3. Первозванский А.А. Математические модели в управлении производством. - М.: Наука, 1975. - 616 с.

4. Корлат А.Н., Кузнецов В.Н., Новиков М.М., Турбин А.Ф. Полумарковские модели восстанавливаемых систем и систем массового обслуживания. - Кишинев: Штиинца, 1991. - 276 с.

5. Obzherin, Yu. E. and A. I. Peschansky (1994) Reliability of unstructured systems with excess time. Cybern. Syst. Anal., Vol. 30, № 6.

6. Obzherin, Yu. E. and A. I. Peschansky (2004) Reliability analysis of system with combined time reserve. Cybern. Syst. Anal., Vol. 40, № 5.

7. Obzherin, Yu. E. and A. V. Skatkov (2010) On the time to failure of systems with large replenishable reserve time. J. Math. Sci., Vol. 57, № 5.

8. Копп В.Я., Обжерин Ю.Е., Песянский А.И. Стохастические модели автоматизированных производственных систем с временным резервированием. - Севастополь: Изд-во СевГТУ, 2000. – 284 с.

УДК 62–50
Л.А. Краснодубец, д-р техн. наук, проф.
А.А. Кабанов, канд. техн. наук, доц.
Севастопольский национальный технический университет, г.Севастополь, Украина
lakrasno@gmail.com, KabanovAleksey@gmail.com

КОНСТРУИРОВАНИЕ ИНФОРМАЦИОННОЙ СТРУКТУРЫ И СИНТЕЗ ПАРАМЕТРОВ ОБРАТНЫХ СВЯЗЕЙ СИСТЕМ ТРАЕКТОРНОГО УПРАВЛЕНИЯ МОБИЛЬНЫМИ РОБОТАМИ

Введение

При траекторном управлении мобильными роботами (МР) возникают две задачи. Первая из них связана с выводом МР на заданную траекторию, вторая – со следованием по этой траектории. В задаче следования траектории наиболее важно, чтобы отклонения текущего положения МР от этой траектории были минимальными. При этом качество следования определяется некоторым критерием, который необходимо минимизировать. В результате минимизации получают законы управления, которые физически реализуется в виде регуляторов. Системы управления с такими регуляторами имеют в своем составе информационно-измерительный канал, структуру которого образуют обратные связи, поставляющие данные о текущем состоянии объекта управления. Эти данные используются для формирования регуляторами управляющих сигналов, подаваемых на исполнительные устройства (приводы), установленные на объектах управления и изменяющие его состояние в соответствии с целью управления. По существу проектирование законов траекторного управления сводится к конструированию информационной структуры и синтезу параметров обратных связей.

Доклад посвящен проектированию законов траек-

торного управления МР на основе методов оптимизации и обратных задач механики. В качестве объектов управления рассматриваются два робота: колесный робот Robotino и гусеничный робот Rover5.

Траекторное управление МР Robotino

Объект управления. Объектом упрвления является МР Robotino, разработанный европейской фирмой Festo и предназначенный для использования в учебном процессе при изучении инженерных дисциплин в области управления и автоматики. Приборная часть Robotino включает процессорный модуль, комплекс сенсоров, видеокамеру, разъемы для подключения дополнительных устройств и модуля доступа к беспроводной сети. Колесная платформа (КП) Robotino имеет круглую форму и оснащена тремя автономными приводами с роликонесущими колесами, установленными под днищем вдоль окружности в точках, разделяющих дуги величиной 120°. Приводы управляются контроллером, обеспечивающим движение КП по трем степеням свободы таким образом, что с места она может начинать продольное движение на плоскости в любом направлении при одновременном вращении вокруг вертикальной оси. Контроллер имеет три управляющих входа, на которые подаются сигналы, соответствующие требуемым линейным скоростям v_x и v_y, а также угловой скорости ω.

Математическая модель движения МР Robotino. Модель движения КП Robotino можно представить в виде:

$$A\ddot{q} = -D\dot{q} + f\ , \quad t = 0 \colon q(0) = q_0, \dot{q}(0) = \dot{q}_0, \tag{1}$$

$$A = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & J \end{bmatrix},\ D = \begin{bmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \eta \end{bmatrix},\ f = [f_x\ \ f_y\ \ h]^T,$$

$$q = [x\ y\ \psi]^T,$$

где m, J – масса и момент инерции робота; μ – коэффициент трения при продольном движении; η – коэффициент трения при вращении; f_x и f_y – проекции силы, приложенной к центру масс; h – вращательный момент; $\omega = \dot{\psi}$ – угловая скорость.

Постановка задачи управления движением МР Robotino. Поставим задачу найти такие управляющие функции – скорости v_x, v_y и ω, обеспечивающие требуемое перемещение на плоскости центра масс КП из начального состояния (1) в заданное конечное состояние $\bar{x}$, $\bar{y}$. Следует отметить, что

Конструкция КП позволяет управлять движением ее центра масс на плоскости за счет применения только двух управляющих функций v_x и v_y при $\omega = 0$. Учитывая это обстоятельство и принимая допущение, что каналы управления по координатам x и y идентичные, можно преобразовать модель (1) к виду

$$\dot{x} + ax = bv_x,\ b = 1;\ a = \mu / m. \qquad (2)$$

В таком случае постановка задачи управления движением центра масс КП Robotino формулируется следующим образом. Для объекта управления, заданного уравнением (2), найти управляющую функцию в форме обратной связи $v_x(t, x, \dot{x})$, которая обеспечит перевод замкнутой системы (2) из начального состояния $t = t_0,\ x(t_0) = x_0,\ \dot{x}(t_0) = 0$ в заданное стационарное состояние равновесия $t > 0,\ x(t) = \bar{x},\ \dot{x}(t) = 0$. При этом потребуем, чтобы процессы $x(t) \to \bar{x},\ \dot{x}(t) \to 0$ в замкнутой системе проходили в малой окрестности эталонных процессов $x^*(t) \to \bar{x},\ \dot{x}^*(t) \to 0$, формируемых эталонной моделью $\dot{x}^* + \alpha_x x^* = \alpha_x \bar{x}$.

Степень близости процессов в управляемой системе и эталонных процессов будем оценивать величиной, опре-

деляемой критерием

$$\Phi(v_x) = 0.5[\dot{x}^* - \dot{x}(v_x)]^2, \qquad (3)$$

который имеет физический смысл нормированной по массе кинетической энергии.

Критерий (3) относится к классу локальных, которые, в отличие от интегральных, вычисляются в процессе функционирования системы управления. При этом искомое управление в виде функции $v_x(t)$ находится из условия минимума (3). В рассматриваемом случае функция управления должна вычисляться бортовым компьютером и, следовательно, она должна иметь ступенчатый вид $v_x(t) = v_x(kT)$ при $kT \le t < kT + T$, где T – период дискретности.

Для вычисления ступенчатой функции $v_x(kT)$ из условия минимума критерия (3) можно воспользоваться рекуррентной процедурой соответствующей простому градиентному методу

$$v_x[(k+1)T] = v_x(kT) - \lambda_x \frac{\partial \Phi(v_x)}{\partial v_x}, \ \lambda_x = const \qquad (4)$$

Производная в правой части (4) вычисляется путем дифференцирования (3) по переменной v_x с учетом того, что она на интервале дискретизации T от времени не зависит. Выражение для искомой производной имеет вид

$$\frac{\partial \Phi(v_x)}{\partial v_x} = (\dot{x}^* - \dot{x})(-1) = -[\alpha_x(\bar{x} - x^*) - \dot{x}].$$

Введем обозначение

$$\theta(t) = [\alpha_x(\bar{x} - x^*) - \dot{x}], \qquad (5)$$

где $\theta(t)$ – информационная функция, которая формируется по данным измерений сигналов обратной связи (по производной) и эталонной модели.

Далее в (5) выполним замену $x^* = x$, что соответствует введению в закон управления обратной связи по координате x вместо сигнала от эталонной модели. Имеем

$$\theta(t) = [\alpha_x(\bar{x} - x) - \dot{x}].$$

Поскольку данные измерений сигналов обратной связи дискретизируются при помощи аналого-цифровых преобразователей, то информационная функция в результате преобразуется к ступенчатой форме $\theta(t) = \theta(kT)$ при $kT \le t < kT + T$.

Подстановка информационной функции в ступенчатой форме в (4) приводит к конечно-разностному уравнению, определяющему дискретный закон управления координатой x:

$$v_x[(k+1)T] = v_x(kT) + \lambda_x\theta(kT).$$

Решение этого уравнения определяет ступенчатую управляющую функция $v_x(kT)$. Аналогичное уравнение получается для нахождения ступенчатой управляющей функции $v_y(kT)$.

Синтез параметров цифрового адаптивного регулятора (коэффициентов при управляемой координате и ее производной, входящих в выражение информационной функции), можно выполнить в соответствии с рекомендациями, приведенными в [1].

Для решения задачи следования заданной траектории на управляющие входы спроектированной системы (по координате x и координате y) вместо констант $\bar{x}$ и $\bar{y}$ следует подавать дискретные последовательности $x_r(t)$ и $y_r(t)$ (временные ряды), определяющие требуемую траекторию заданного движения КП в определенной системе координат. Эти последовательности могут быть решениями конечно-разностных уравнений или задаваться в виде числовых массивов. Следует отметить, что при задании на-

чальных условий движения МР, которые не соответствуют расположению его центра масс на назначенной траектории, сконструированный закон управления обеспечивает последовательное решение двух задач траекторного управления: выведения и следования.

Траекторное управление гусеничным МР Rover5

Объект управления. Объектом исследования является МР, построенный на базе гусеничного шасси Rover5 производства фирмы DAGU. Силовая часть шасси включает в себя два электропривода (по оному на каждую гусеницу) на базе двигателей постоянного тока . В качестве устройства управления используется контроллер Orangutan SVP-324p фирмы Pololu Corporation на базе микроконтроллера ATmega324p от Atmel, содержащий два драйвера двигателей с током до 4 А на один канал.

Математическая модель движения МР Rover5. Математическая модель МР при движений в горизонтальной плоскости с учетом динамики исполнительных двигателей (считаем, что оба двигателя являются полностью идентичными) имеют вид [2]:

$$
\begin{aligned}
&\dot{x} = v_x = v\cos\psi = 0.5(v_1 + v_2)\cos\psi,\\
&\dot{y} = v_x = v\sin\psi = 0.5(v_1 + v_2)\sin\psi,\\
&\dot{\psi} = \omega = 0.5d^{-1}(v_2 - v_1),\\
&m\dot{v} = c_m \cdot n \cdot r^{-1}(I_1 + I_2),\\
&J\dot{\omega} = c_m \cdot n \cdot d(I_2 - I_1),\\
&L\dot{I}_1 = -R_m \cdot I_1 - c_e \cdot n \cdot r^{-1}(v + d \cdot \omega) + U_1,\\
&L\dot{I}_2 = -R_m \cdot I_2 - c_e \cdot n \cdot r^{-1}(v - d \cdot \omega) + U_2.
\end{aligned}
\qquad (6)
$$

где v_1, v_2, v – линейные скорости левой, правой гусениц и самого МР соответственно; d – расстояние от оси МР до гусеницы; n – передаточное отношение редуктора; r – радиус ведущего колеса; c_m, c_e – механическая и электриче-

ская постоянные двигателя; I_i – ток в якорной обмотке двигателя; R_m, L – сопротивление и индуктивность якорной обмотки двигателя; U_i – напряжение в цепи якоря двигателя.

Постановка задачи управления движением МР Rover5. Программное движение робота Rover5 по заданной траектории определяется координатами некоторой произвольной точки *M* робота $x_r\ y_r$. Проекции скорости точки *M* на неподвижные оси координат являются непрерывными функциями времени $v_{xr} = \dot{x}_r$, $v_{yr} = \dot{y}_r$.

Особенностью задачи траекторного управления МР Rover5, который представляет собой неголономную систему, являются ограничения, возникающие при задании программных траекторий и определяемые неголономной связью вида [2]

$$-v_x \sin\psi + v_y \cos\psi = 0.$$

Наличие данной связи не позволяет произвольно задавать угловую координату робота ψ, которая в данной ситуации должна быть решением дифференциального уравнения [2]

$$\omega_r = \dot{\psi}_r = \frac{-v_{xr}\sin\psi_r + v_{yr}\cos\psi_r}{b_M \cos\gamma},\ \psi_r(0) = \psi_0, \qquad (7)$$

где $b_M \cos\gamma$ и $b_M \sin\gamma$ – постоянные координаты точки *M* в подвижной системе координат, связанной с роботом. Интегрируя уравнение (7), находим закон изменения заданного курсового угла ψ_r.

Степень отклонения состояния управляемой системы от заданных параметров движения определяет функционал качества

$$\rho = \frac{1}{2}\int_0^{\infty}(q_1^2 v + q_2^2 \omega + r_1^2 U_1 + r_2^2 U_2)dt, \quad q_1, q_2, \geq 0, r_1, r_2 > 0,$$

где q_1, q_2, r_1, r_2 – коэффициенты штрафов на состояние и управление.

Решение задачи минимизации критерия (7) на движениях системы (6) приводит к линейному закону управления:

$$u = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} K_{11} v + K_{12} \omega \\ K_{21} v + K_{22} \omega \end{pmatrix},$$

где $K_{i,j}$ – постоянные коэффициенты.

Синтез параметров регулятора $K_{i,j}$ можно выполнить в соответствии с рекомендациями, приведенными в [2].

Список литературы

1. Краснодубец Л. А., Крамарь В.А. Проектирование адаптивных регуляторов для мехатронных систем // Мехатроника, автоматизация, управление, 2012. № 1. – С. 8.–12.

2. Кабанов А.А. Система робастного субоптимального управления движением мобильного робота // Мехатроника, автоматизация, управление, 2013. № 4. – С. 14–19.

УДК 681.3: 004-52
М.П. Фархадов, д-р. техн. наук, зав. лаб.
В.А. Вертлиб, канд. техн. наук, вед. научн. сотр.
Н.В. Петухова, ст. научн. сотр.
З.П. Мясоедова, научн. сотр.
Институт проблем управления им В.А.Трапезникова РАН, г. Москва, Россия
mais@ipu.ru, vertlibipu@.ru, nvpet@ipu.ru, mzinap@mail.ru

ВОПРОСЫ СОЗДАНИЯ РАСПРЕДЕЛЕННЫХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ НАСЕЛЕНИЯ С СОВРЕМЕННЫМИ РЕЧЕВЫМИ ИНТЕРФЕЙСНЫМИ ТЕХНОЛОГИЯМИ

Введение

Актуальность исследований определяется наличием к настоящему времени большого количества разнородных и изолированных систем, что объясняется интенсивным развитием компьютерных, информационных и коммуникационных технологий в последнее десятилетие. Этот процесс, однако, имеет и свои негативные стороны, состоящие в том, что информация многократно дублируется, ее достоверность снижается, защита данных не отвечает требованиям к безопасности, а неравенство в доступе к данным растет. Решение этих проблем лежит на путях организации сетевых структур нового типа, отвечающих современным требованиям по реактивности, защищенности, доступности и эффективности.

Автоматизированные системы массового обслуживания населения

Развитие технологий, средств вычислительной техники и связи привело к практическому внедрению в разных областях хозяйственной деятельности автоматизированных систем массового обслуживания (АСМО), основным на-

значением которых является предоставления населению информации и услуг.

В таких системах обеспечивается реально-временной доступ большого количества удаленных пользователей к рассредоточенным по территории базам данных, содержащим некие ресурсы. С точки зрения инфраструктурной организации автоматизированная система этого класса, как и любая современная распределенная система, неизбежно содержит три компонента: систему баз данных, размещаемых в центрах обработки данных (ЦОД), информационно-коммуникационную сеть (ИКТ) и периферийную (абонентскую, терминальную) технику. Система баз данных содержит данные о всех ресурсах, клиентах-поставщиках этих ресурсов, клиентах-потребителях ресурсов и другие сведения, необходимые для контроля и управления состоянием и предоставлением ресурсов, включая сбор статистики, защиту от злоупотреблений и ошибок, ранжирование доступа и т.п. Периферийная техника здесь обозначает все многообразие персональных и коллективных средств доступа, обеспечивающих абоненту формирование сообщений для системы, поддержание диалога с системой через информационно-коммуникационную сеть связи на уровне системной транзакционной технологии обмена сообщениями. Сеть связи (ИКТ) служит для переноса информационных сообщений от точки доступа абонента к сети до точки входа в систему банков данных и обратно. В работах [1-4] рассматриваются вопросы построения автоматизированных систем массового обслуживания коммерческого и некоммерческого назначения, ориентированных на применение в областях «электронной коммерции», государственных учреждениях, транспортных и туристических системах обслуживания населения. В [5] описывается принципы работы классических автоматизированных систем массового обслуживания населения, их функционально-технологического наполнение как части любой автоматизированной системы управления, ориентированной час-

тично или полностью на предоставление услуг населению.

Автоматизированные системы массового обслуживания населения с самого начала своего существования имели в своем составе развитые сетевые структуры, которые являлись важнейшей их составляющей, даже на этапе автономного существования этих систем и их моноцентральной архитектуры. Эволюция глобальных сетей общего пользования, развитие Интернет - технологий и конкуренция на рынке провайдеров Интернет услуг обеспечили условия для создания и функционирования нового поколения автоматизированных систем массового обслуживания разной ориентации [6]. Простой и дешевый доступ к сетевым услугам и информационным базам сети Интернет и возможность легкого включения в Интернет собственных информационных баз привели к внедрению в рамках этой сети систем массового обслуживания различной направленности и различного масштаба. При этом три главных фактора стимулируют интерес к сетевым информационным технологиям, обеспечивающим создание и функционирование систем для торговых, консалтинговых, справочных, игровых и т.п. систем массового обслуживания, действующих в сети Интернет.

- Первый фактор связан с достижением высоких скоростей передачи информации в сети, обеспечивающих обмен в реальном масштабе времени различного типа информации, включая аудио- и видео- данные, причем с качеством и стоимостью приемлемыми для конечных пользователей. При этом комплексность предоставляемых услуг по сути уникальна, т.к. фактически интегрируется в одном «модуле» все связевое и информационное прикладное обслуживание.
- Второй фактор - это возможность приобретения технических средств доступа к услугам сети и данным функционирующих в ней систем широкого круга пользователей, не являющихся техническими специалистами в областях информационных технологий.

- Третий фактор (возможно главный) обусловлен существенным ростом информационно-компьютерной грамотности населения, которое освоило компьютерные и сетевые технологии взаимодействия, в том числе и мобильные, отчетливо понимают сегодняшние и перспективные возможности использования этих технологий в повседневной жизни. За растущие удобства и функциональность люди готовы платить, вследствие чего растут объемы и характер предоставляемых услуг, повышается номенклатура и качество технических средств, расширяются и развиваются системотехнические и прикладные методы организации и эксплуатации коммуникационно-компьютерных образований.

Анализ показал, что функциональное объединение компьютерных и коммуникационных средств в глобальные вычислительные сети привело к синергетическому эффекту и породило множество концептуальных идей в различных сферах жизни: экономике, политике, военном деле, управлении социальными и общественными образованиями и т.д. Реализация компьютерных возможностей в мобильных телефонах, планшетах, ноутбуках и т.п. индивидуальных приборах, в принципе ориентированных на постоянное сопровождение индивидуума и обеспечивающих ему доступ к сети, ее базам данных, отдельным ее абонентам, независимо от их места нахождения и в реальном времени – это неизбежно ведет к новому витку развития информационно-коммуникационных технологий.

Количественный и функциональный рост компонентов вычислительных сетей, повышение сложностей защиты и эксплуатации, постоянные корректировки и модернизации программного и технического обеспечения, вызвали необходимость в разработке особых методов организации и обслуживания систем, улучшающих экономические и качественные характеристики глобальных сетей и вычислительных средств. Многие методы обслуживания и эксплуатации таких объектов уже широко практикуются в бизнесе.

Структура для осуществления взаимодействия систем массового обслуживания с использованием речевых и телекоммуникационных технологий

Автоматизированная обслуживающая функциональная система. Обычно обслу-живающая система содержит две подсистемы: «автоматизированную информационную» и «исполнительную материальную».

Автоматизированная информационная обслужи-вающая система (подсистема) содержит информацию о некотором материальном ресурсе (например, места на самолеты, места в гостинице, время бракосочетания, время посещения врача и т.п.) и спроектирована как аппаратно-программное средство, предназначенное для обработки и предоставления этого ресурса потребителям по их запросам (требованиям), для обновления ресурса поставщиком ресурса и управления процессом обслуживания по командам администратора системы. Для существования обслуживающей информационной системы достаточно наличия ресурса, комплекса систем автоматизированного управления, пользователей системы (поставщика «продавца» и потребителя ресурса «покупателя») [4].

Исполнительная материальная система (подсистема) обеспечивает потребителю, т.е. каждому гражданину, обслужившемуся в информационной системе и получивший соответствующий ресурс, «материальное обслуживание» (например, фактическую посадку на выбранный авиарейс, проживание в соответствующем номере гостиницы, посещение в указанное время ЗАГСа для получения государственной услуги бракосочетания, врачебного обслуживания во время и в обусловленной поликлиники и т.п.).

Сетевая обслуживающая система. Предполагается, что в глобальной сети Интернет имеются серверы (порталы, сайты) государственных и частных юридических и физических лиц, имеющих некоторый ресурс (или совокупность ресурсов). Аппаратно-

программные средства серверов ориентированы на сетевую работу в Интернете, т.е. имеют собственный адрес в глобальной сети и выполняют все необходимые сетевые протокольные процедуры. На сайте-«витрине» системы представлены необходимые правила работы пользователей, клиентов, «гостей» и т.п.

Совокупность аппаратно-программных вычислительных и коммуникационных средств образуют центр обработки данных (ЦОД) в качестве структурной единицы функциональной обслуживающей системы, находящейся под единым административным управлением. Каждый центр обработки данных является функционально определенной компонентой сети, ориентированной на обслуживание известных видов ресурсов. Центры обработки данных могут иметь различные по типу и вычислительной мощности технические средства, различные по номенклатуре и объемам ресурсы, включая информацию о клиентах, и вычислительные мощности для обслуживания заявок потребителей, различные категории обслуживаемых потребителей. Объединяющим и системообразующим признаком АСМО является техническая и технологическая способность ЦОД взаимодействовать через глобальную сетевую среду Интернет друг с другом и с пользователями в реальном масштабе времени.

Под термином «сетевая обслуживающая система» понимается множество центров обработки данных (ЦОД), взаимодействующих через телекоммуникационную среду Интернета между собой и терминальными средствами пользователей. Телекоммуникационная среда Интернета – множество взаимодействующих сетей передачи данных, в совокупности обеспечивающие реализацию разнообразных телекоммуникационных услуг с заданными параметрами качества. Перечень используемых при этом сетей включает локальные корпоративные и глобальные сети и сети общего пользования. Важна согласуемость сетевых протоколов. Терминальные средства потребителей – разнообразные

персональные компьютеры и интеллектуальные дисплейные комплексы, мобильные телефоны и другие интерфейсные средства на базе мобильной персональной техники, способные обеспечить пользователям надлежащее их статусу взаимодействие через Интернет с ЦОД для доступа к соответствующим ресурсам. По существу, терминальные средства в совокупности с оконечным портом сети образуют точку доступа в сеть (в функциональном смысле). Такой точкой доступа может быть порт локальной сети, имеющей выход в Интернет, модемный канал на коммутируемый вход провайдера Интернет, групповые средства заведений типа Интернет-кафе и т.п.

Критерии качества. Вопрос распределения услуг имеет два конкурирующих мотива: мотив «продавца», т.е. структуры предоставляющей ресурсы, и мотив «покупателя», т.е. структуры, потребляющей ресурсы. В рассматриваемом случае потребляющей стороной предполагается «население», т.е. множество отдельных граждан, не связанных друг с другом каким-либо служебным интересом с функционированием обслуживающей системы. Эти граждане заинтересованы в точной и четко регламентированной работе обслуживающей системы, в частности, гарантированном времени обслуживания, удобстве доступа, четких формулировках каждой предоставляемой услуги и т.п., а также полного соответствия материального обеспечения с провозглашенной в информационном обслуживании услуге. В конкретных случаях как информационное обслуживание может потребовать работу с несколькими ЦОД (базами данных), так и материальное обслуживание может быть связано с несколькими службами обеспечения (у которых также могут использоваться вычислительные средства) и отдельными функционерами этих служб.

Критерии качества работы указанных подсистем могут не совпадать.

Обслуживающая система содержит некоторый ресурс и спроектирована как аппаратно-программное средст-

во, предназначенное для предоставления этого ресурса потребителям по их запросам (требованиям), для обновления ресурса поставщиком ресурса и управления процессом обслуживания по командам администратора системы. Для существования обслуживающей системы достаточно наличия ресурса, комплекса систем автоматизированного управления, пользователей системы (поставщика и потребителя ресурса) [4].

Сетевая обслуживающая система. Предполагается, что в глобальной сети Интернет имеются серверы (порталы, сайты) государственных и частных юридических и физических лиц, имеющих некоторый ресурс (или совокупность ресурсов). Аппаратно-программные средства серверов ориентированы на сетевую работу в Интернете, т.е. имеют собственный адрес в глобальной сети и выполняют все необходимые сетевые протокольные процедуры. На сайте-«витрине» системы представлены необходимые правила работы пользователей, клиентов, «гостей» и т.п.

Совокупность аппаратно-программных вычислительных и коммуникационных сред-ств образуют центр обработки данных (ЦОД) в качестве структурной единицы функциональной обслуживающей системы, находящейся под единым административным управлением. Каждый центр обработки данных является функционально определенной компонентой сети, ориентированной на обслуживание известных видов ресурсов. Центры обработки данных могут иметь различные по типу и вычислительной мощности технические средства, различные по номенклатуре и объемам ресурсы, включая информацию о клиентах, и вычислительные мощности для обслуживания заявок потребителей, различные категории обслуживаемых потребителей. Объединяющим и системообразующим признаком АСМО является техническая и технологическая способность ЦОД взаимодействовать через глобальную сетевую среду Интернет друг с другом и с пользователями в реаль-

ном масштабе времени.

Под термином «сетевая обслуживающая система» понимается множество центров обработки данных (ЦОД), взаимодействующих через телекоммуникационную среду Интернета между собой и терминальными средствами пользователей. Телекоммуникационная среда Интернета – множество взаимодействующих сетей передачи данных, в совокупности обеспечивающие реализацию разнообразных телекоммуникационных услуг с заданными параметрами качества. Перечень используемых при этом сетей включает локальные корпоративные и глобальные сети и сети общего пользования. Важна согласуемость сетевых протоколов. Терминальные средства потребителей – разнообразные персональные компьютеры и интеллектуальные дисплейные комплексы, мобильные телефоны и другие интерфейсные средства на базе мобильной персональной техники, способные обеспечить пользователям надлежащее их статусу взаимодействие через Интернет с ЦОД для доступа к соответствующим ресурсам. По существу, терминальные средства в совокупности с оконечным портом сети образуют точку доступа в сеть (в функциональном смысле). Такой точкой доступа может быть порт локальной сети, имеющей выход в Интернет, модемный канал на коммутируемый вход провайдера Интернет, групповые средства заведений типа Интернет-кафе и т.п.

Вопросы сопряжения разноплановых систем. В техническом аспекте реализация этой задачи связана с разработкой интерфейсных решений, как для сопряжения разнородных систем, так и подсистем, объединяемых в единую систему, но разрабатываемых независимо и разновременно. Вопросы сопряжения унаследованных систем в единую работающую автоматизированную систему в общем случае уникальны и в лучшем случае определяются для каждой пары сопрягаемых комплексов. Решение шлюзовых задач на автоматической базе и согласование протоколов почти всегда требует коррекции функционирования

в сопрягаемых системах.

Внедрение таких объединяемых систем связано с большими проблемами в плане их взаимодействия, т.к. часто сопрягаемые компоненты не обладают достаточной функциональной гибкостью и универсальностью. Наилучшие результаты при сопряжении разнотипных систем могут быть достигнуты при использовании концепции распределительных структур, оперирующих с отдельными системами, содержащими разноплановые ресурсы, в едином информационном пространстве.

Примером является ситуация в туристическом бизнесе, когда одновременно заказывается комплекс туристических услуг: транспорт – гостиница – автомобиль – экскурсии и др. Для решения этих задач появились распределительные системы, которые предоставляют услуги аутсорсинга многим фирмам, имеющим какой-нибудь ресурс, и подключают тысячи и десятки тысяч терминалов «агентов», рассредоточенных по всему миру. Одной из первых была авиационная система резервирования SABRE, которая превратилась в глобальную распределительную систему (Global Distribution System, GDS), предназначенную для оказания путешественникам информационных услуг, резервирования мест и обработки транзакций. Система связывала более 30 тыс. агентов туристических бюро и 3 млн. интерактивных клиентов с 400 авиакомпаниями, 50 фирмами по прокату автомобилей, 35 тысячами отелей и тысячами железнодорожных и туристических компаний.

В России системы резервирования разделились на системы резервирования авиакомпаний, которые называют инвенторными системами, и распределительные системы. Посредством инвенторных систем осуществляется продажа мест на рейсах авиакомпаний: через собственную терминальную сеть этих систем в основном агентами авиакомпаний, но также и через терминальные сети других систем резервирования. Инвенторная система может целиком принадлежать отдельной авиакомпании, а может представ-

лять собой систему коллективного использования и предоставлять услуги нескольким перевозчикам.

Распределительные системы, в отличие от систем резервирования авиакомпаний, не связаны с размещением и управлением ресурсами авиакомпаний. Их исключительной задачей является продажа авиаперевозок нейтральным агентам. Чтобы в процессе обслуживания пассажира осуществить резервирование места на рейсе какой-либо авиакомпании, из распределительных систем реализован доступ в инвенторные системы. При этом весь диалог с агентом по подбору маршрута, тарификации и продаже перевозки осуществляется на языке данной распределительной системы независимо от типа инвенторной системы, что является одним из главных преимуществ распределительных систем. Наиболее крупные распределительные системы определяют как Глобальные (ГРС).

Сегодня, чтобы выйти на тот или иной агентский рынок, авиакомпании достаточно подключить свою инвенторную систему к ГРС, обслуживающей этот рынок. При этом благодаря стандартным интерфейсам и протоколам, связывающим эти системы, задача подключения к ГРС не составляет технических проблем. Она сводится в основном к подписанию коммерческого соглашения между ГРС и авиакомпанией о распределении ресурсов мест последней среди агентов, использующих терминальную сеть данной ГРС (так называемое «Соглашение об участии перевозчика»).

ГРС вышли за пределы авиационной отрасли и предлагают своим пользователям комплексное обслуживание, которое включает бронирование в гостиницах, аренду автомобилей, продажу железнодорожных билетов. Таким образом, термин «Глобальная» можно интерпретировать в двух направлениях:

- во-первых, обеспечение доступа агентов к ресурсам мест практически всех основных авиаперевозчиков. Благодаря этому агент получает возможность, используя

терминал данной системы, формировать авиамаршруты любой сложности, охватывающие географию всего мира;

- во-вторых, предоставление пассажиру комплексного обслуживания, включая бронирование гостиниц, ренту автомобилей и другой туристический сервис.

Важно отметить, что ГРС обладают свойствами, характерными для сетецентрических и клиентоцентрических систем: осведомленностью о текущей ситуации и статусе иерархически распределенных структур, возможностью доступа к локальным и удаленным ресурсам, ориентацией в информационном пространстве, интероперабельностью, обеспечивающей доступ к сервисам разнородных клиентов, динамически обновляемой распределенной базой данных.

Важной тенденцией в современном мире является развитие компьютерных возможностей в мобильных телефонах, планшетах, ноутбуках и т.п. индивидуальных приборах, в принципе ориентированных на постоянное сопровождение индивидуума и обеспечивающих ему доступ к сети отдельным ее абонентам, независимо от их места нахождения и в реальном времени. Это неизбежно приводит к необходимости развития новых интерфейсных технологий и к новому витку развития информационно-коммуникационных систем обслуживания.

Основные показатели качества работы систем массового обслуживания. Исходя из основной задачи АСМО – обеспечения эффективного взаимодействия Хранилища ресурсов и Потребителя ресурса (Поставщика ресурса) очевидно, что основные показатели системы как функционера именно массового обслуживания должны определяться временными характеристиками: временем обслуживания, временем ожидания обслуживания, временем реакции. Эти критерии должны быть точно определены и регламентированы с учетом реальных условий и производительности технических средств. Автоматическое измерение и обработка этих характеристик может достаточно

легко выполняться, так что возможно полное исключение субъективного фактора в оценке качества работы системы. На эти параметры, по существу, работают все (или почти все) другие характеристики, свойственные большим компьютерным комплексам и характеризующие систему с позиций технического функционирования ее как по составляющим компонентам, так и в целом: надежность, достоверность, безопасность, связность сети, времена простоя и ремонта и т.п.

Для АСМО характерны случайные потоки входных заявок, случайные многофазные процессы обслуживания заявок, стохастичные процессы управления ресурсами. Это делает возможным, а часто необходимым, применение моделей теории массового обслуживания при расчетах технических параметров системных комплексов, например, числа и производительности трактов передачи и обработки данных, вычислительных средств, объемов памяти, при анализе бизнес-моделей и расчетах стратегических вариантов распределения ресурсов, маркетинговых процессов и т.п.

Принципы и решения для построения глобальной системы массового обслуживания населения и модель идеальной системы. Разработка архитектурных решений и технологической модели для идеальной системы

Как было показано, наиболее сложными вопросами создания современных систем массового обслуживания населения являются вопросы сопряжения и взаимодействия раз-нородных, распределенных в сетевом пространстве систем, и вопросы, связанные с раз-работкой интерфейсов пользователей, имеющих разнообразные оконечные устройства и обладающих разными по уровню умениями и навыками работать с современным оконечным оборудованием и технологиями. Идеальная система

должна, как и прежде, обеспечивать вы-полнение требований, предъявляемых к функционированию систем массового обслу-жи-вания, но также отвечать и двум новым главным условиям: обеспечивать взаимодействие всех систем, чьи ресурсы запрашивают клиенты, и предоставлять клиентам возможность запроса услуг с разных оконечных устройств. На рисунке 1 показана принципиальная схема архитектурного решения для системы массового обслуживания, которая позволяет раз-но-родным клиентам получать услуги и данные из различных инвентарных систем.

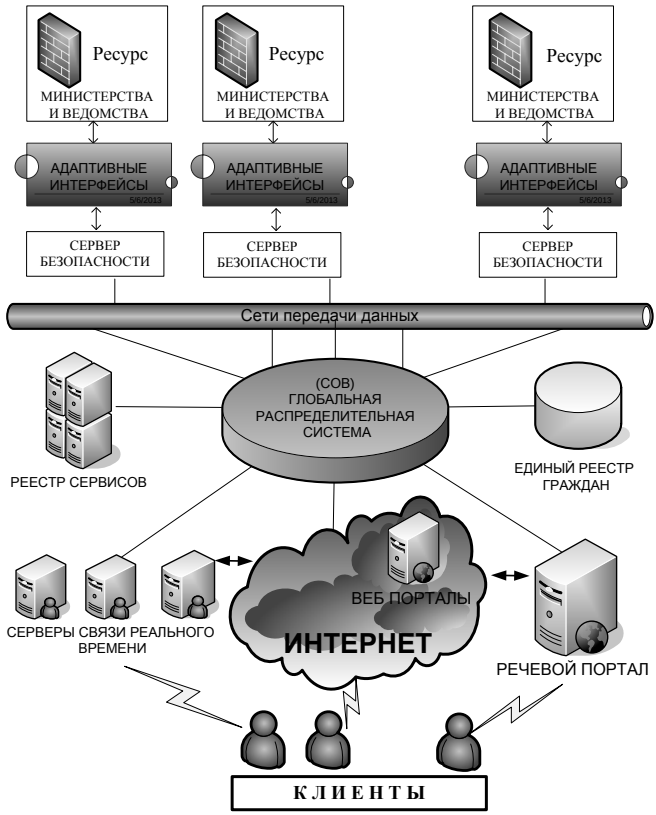


Рисунок 1. Архитектура идеальной системы массового обслуживания

Взаимодействие инвенторных систем осуществляется через систему организации взаимодействия – глобальную распределительную систему (СОВ - ГРС). Принципиальными задачами СОВ являются следующие:

- представление клиентам в едином формате полной и актуальной информации об услугах и состоянии ресурсов участвующих в обслуживании систем;
- представление единого интерфейса для взаимодействия между системами для совершения транзакций;
- подготовка данных для транзакций;
- регистрация транзакций,
- отслеживание транзакций и информирование о результатах клиентов и системы;
- информационная поддержка взаиморасчетов.

Если рассматривать такое решение применительно к организации систем электронного правительства, то СОВ будет выполнять следующие функции:

- единый интерфейс взаимодействия граждан со сторонними системами - участниками СОВ;
- регистрация граждан в системе и представление им электронных средств идентификации и аутентификации личности;
- оперативное информирование граждан о наличии / отсутствии персональных данных о них во всех СУ;
- возможность предусмотренных законом действий граждан с персональными данными во всех СУ;
- информирование граждан о запросах и доступе к их персональным данным в СУ и целях этих запросов;
- информирование граждан о состоянии его заявок на услуги;
- подтверждение сведений о гражданине по запросу юридического лица;
- оформление заявок на изменение,

подтверждение личных прав гражданина в соответствии с законами РФ;

- оформление сделок и действий, изменяющих права нескольких лиц;
- оформление доверенностей на действия;
- информационная поддержка работы адвокатуры, нотариальных контор, кадровых служб предприятий, служб занятости и других организаций, работающих с населением и их подразделений;
- запись на прием.

Таким образом, СОВ в идеальной распределенной системе массового обслуживания по существу представляет собой внешнюю для участвующих систем интеграционную шину для обмена данными и регистрации транзакций. Для реализации этой шины необходима единая модель данных и алгоритмов представления услуг, единый интерфейс к базам данных и единое описание услуг и алгоритмов. Информация о транзакциях фиксируется в единой технологической базе данных, позволяющей реализовать алгоритмы взаимодействия различных систем в процессе представления услуг, а также схему взаиморасчетов, позволяющую распределить средства, поступившие от клиентов и коммерческих организаций, подключенных к системе в соответствии с произведенными транзакциями.

Интерфейсные решения на основе речевых и мультимедийных технологий. Речевые и мультимедийные технологии являются в настоящее время наиболее востребованными и эффективными для расширения возможностей пользователей при взаимодействии с информационными и сервисными системами [7, 8, 9]. Однако проведенные исследования и опыт разработки подобных интерфейсов показали, что создание удобного и эффективного интерфейса является непростой задачей. В частности, было показано, что речевой интерфейс отличается от графического наличием более серьезных ограничений по взаимодействию как с машинной стороны, так и со стороны чело-

века. В связи с этим разработка принципов и методов создания эффективного речевого интерфейса включала, в первую очередь, исследование проблематики каждой из сторон интерфейса: распознавателей с одной стороны и человека с другой. Было выполнено экспериментальные исследования свойств распознавателей. Для исследования характеристик распознавателей русского языка и выявления зависимостей создана система, обеспечивающая доступ к реальным распознавателям через телефон. Система построена на компьютере Pentium 4 (операционная система Windows 2000/XP), дополненном многоканальной интерфейсной платой Dialogic, к которой подключены телефонные линии. Использовались системы распознавания речи разработки компаний Nuance Communications и Philips.

Для выполнения экспериментов был разработан речевой диалог, созданы речевые блоки и записаны аудио файлы для озвучивания текста.

Тестирование производилось различными людьми, разного возраста, пола и национальности, что позволило представить достаточно полный набор вариантов произношения.

Выбор грамматик (грамматика – текстовый файл, описывающий множество слов и фраз, которые могут быть произнесены клиентом на данном шаге диалога) осуществлялся таким образом, чтобы исследовать с одной стороны специфичные для систем массового обслуживания грамматики (цифры, числа, даты, месяцы, пин коды, города, улицы, телефонные номера, названия валют и др.), и с другой стороны охватить все наиболее интересные для исследования аспекты проблем распознавания. Грамматики отличались размером, сложностью логики, длиной произносимых фраз, длиной слов.

Помимо лабораторных данных, использовались сохраняемые записи реальных вызовов, поступающих в работающую систему «Автосекретарь», в которой применено распознавание речи. Для исследования были взяты двухне-

дельные фрагменты архива, содержащие около 3000 логов клиентов с системой каждый. Звуковым файлам с записями речи клиентов было поставлено в соответствие несколько информационных меток: смысловое значение, качество записанного звука, половая принадлежность клиента, темп и громкость речи. Смысловое значение использовалось для определения правильности распознавания имен. Грамматика системы «Автосекретарь» содержит около 300 слов, представляющих собой имена, фамилии, отчества. Исследование проводилось путём многократной обработка звуковых файлов системой распознавания при различных значениях параметров распознавателя с последующим запоминанием результатов экспериментов.

Изучались также свойства человека в контексте речевого интерфейса. Главной проблемой унимодального речевого интерфейса является проблема невидимости: зрение не участвует в процессе взаимодействия с машинной стороной, и восприятие вопросов и ответной информации производится клиентом только лишь на слух. Исключение зрения из процесса взаимодействия с машиной приводит к значительному возрастанию когнитивной нагрузки на человека. Показано, что наибольшее внимание в процессе дизайна речевого интерфейса требуется уделять следующим составляющим когнитивной нагрузки: нагрузке на память, удержанию внимания и понятийной сложности.

На основании полученных результатов разработаны методы решения проблем и сформулированы правила проектирования эффективных интерфейсов для речевого взаимодействия клиентов с системами массового обслуживания.

Использование речевых и мультимедийных технологий для организации пользовательского интерфейса было реализовано в системе «Звуковой тренажер» в онлайн доступом к ее ресурсам.

Эта система представляет собой *звуковой тренажер для обучения людей с нарушениями слуха правильному про-*

изношению звуков русского языка. Принцип действия тренажера состоит в спектральном преобразовании входного аудио сигнала, выявлении его характеристик, сравнении с эталоном и в возврате и представлении результата. В нашей стране по данным Всероссийского общества глухих сотни тысяч людей страдают нарушениями слуха. Включение их в трудовую и общественную жизнь тесно связано с обучением их устной вербальной речи. Как правило, артикуляционный аппарат этих людей не нарушен, но пользоваться им они не умеют без специального обучения. Целью создаваемого компьютерного тренажера являлась не замена сурдопедагога, а предоставление пользователям возможностей для дополнительных самостоятельных занятий. Такие возможности особенно важны при обучении детей с пониженным слухом или с нарушениями артикуляции, поскольку к процессу обучения детей могут подключаться родители, да и сами дети получат возможность отрабатывать артикуляцию дома, а не только в кабинете педагога.

Размещение подобных программ в интернете и предоставление онлайн доступа к ним значительно увеличивает число людей, получающих возможность тренировать звукопроизношение с домашних компьютеров.

Для решения поставленной задачи было принято решение использовать компьютерное распознавание речи в качестве анализатора поступающего от пользователя аудио сигнала. Ядро распознавателя и соответствующие программы обработки и управления располагаются в центральной части обрабатывающего комплекса, размещаемого на сервере. Пользователи получают на экран необходимые изображения, инструкции и подсказки. Специальная справочная информация дает пользователям общие сведения о звуках и подробное описание способа произнесения каждого отрабатываемого звука. Разработанная в рамках данного проекта система представляет собой тренажер для обучения и отработки звуков русского языка с онлайн доступом через интернет.

Большое внимание было уделено привлекательности интерфейса. Как показали проведенные исследования, наиболее важную информацию неслышащие люди получают, наблюдая движения губ. В связи с этим было принято решение поместить на экран два окна, где демонстрируется движение губ при произнесении выбранного звука (вид анфас и в профиль). Используется также окно, где можно наблюдать анимационное изображение работы артикуляционного аппарата для выбранного звука. Приводится текстовое описание работы речевого тракта и предоставлена возможность демонстрации звука путем нажатия соответствующей кнопки. Пользователю также обеспечена возможность контроля движения своих губ при произнесении им выбранного звука, если к его компьютеру подключена веб-камера. Для обеспечения обратной связи предусмотрена визуализация звука: на экран выводится изображение импульсно-кодовой модуляции эталонного и произнесенного звука.

Оценка качества произнесения звука производится как в числовой форме по 100-балльной шкале, так и графическим методом путем вывода на экран смайлика (используется условная 5-балльная шкала).

Заключение

Показана дистрибутивно-инвенторная архитектура для создания распределенных систем обслуживания, в которой унифицированное внутрисетевое взаимодействие инвенторных систем осуществляется через систему организации взаимодействия.

Показана необходимость разработки новых интерфейсных решений для организации взаимодействия пользователей с системами. Выявлены проблемы, возникающие при разработке удобного и эффективного речевого интерфейса, и предложены обоснованные методы их решения.

Разработан ряд действующих прикладных систем

для предоставления гражданам информации и услуг, в которых использованы полученные теоретические результаты.

Список литературы

1. Силаев В.Н., Вертлиб В.А., Маргулис Д.С. Диалоговая связь в телеавтоматических системах массового обслуживания. Под ред. В.Н.Силаева. – М.: Энергия, 1977.
2. Билик Р.В., Вертлиб В.А. Автоматизированные системы массового обслуживания в Интернете. – М.: МАКСПресс, 2001, – 31 с.
3. Билик Р.В., Вертлиб В.А. Системы электронной коммерции в Internet - как эволюция автоматизированных систем массового обслуживания // Труды ИПУ им. В.А. Трапезникова РАН, Т. IX. – М.: ИПУ РАН, 2000.
4. Билик Р.В., Вертлиб В.А., Гуденко А.А. Методология автоматизированных систем массового обслуживания – база современной сетевой экономики // Проблемы управления. 2006. № 2. - С. 8-13.
5. Вертлиб В.А., Фархадов М.П., Петухова Н.В. «Электронное Государство» как автоматизированная система массового обслуживания населения / ИПУ им. В.А.Трапезникова РАН. – М.: Макс Пресс, 2008. - 147 с.
6. Жожикашвили В.А., Вертлиб В.А., Ребортович Б.И., Петухова Н.В., Фархадов П.П. «Автоматизированные системы управления процессами массового обслуживания нового поколения» // Труды ИПУ им. В.А.Трапезникова РАН, Т. IV. – М.: ИПУ РАН, 1999.
7. Фархадов М.П. Распознавание речи в системах массового обслуживания населения // Труды СПИИРАН. 2011. Вып. 19. - С. 65-86.
8. Билик Р.В., Жожикашвили В.А., Петухова Н.В., Фархадов М.П. Анализ речевого интерфейса в интерактивных сервисных системах. I // Автоматика и телемеханика. 2009. № 2. - С. 80-89.
9. Билик Р.В., Жожикашвили В.А., Петухова Н.В., Фарха-

дов М.П. Анализ речевого интерфейса в интерактивных сервисных системах. II // Автоматика и телемеханика. 2009. № 3. - С. 97-113.

УДК 681.3: 004-52
В.А. Смирнов[1]**, ген. дир.**
М.П. Фархадов[2]**, д-р техн. наук, зав. лаб.**
[1]*ООО «Спич-Драйв», г. Санкт-Петербург*
[2]*Институт проблем управления им В.А. Трапезникова РАН, г. Москва, Россия*
speechdrive@mail.ru, mais@ipu.ru

СИСТЕМА АНАЛИЗА НЕСТРУКТУРИРОВАННОЙ РЕЧЕВОЙ ИНФОРМАЦИИ В СЕТИ ИНТЕРНЕТ

Введение

В настоящее время наблюдается повышенный интерес к системам, помогающим государственным организациям и коммерческим компаниям решать бизнес-задачи, осуществлять аналитику через сеть Интернет, без инсталляции какого-либо программного обеспечения в инфраструктуре. Подобные системы популярны среди клиентов, потому что позволяют, с одной стороны, не приобретать дорогостоящие лицензии, а с другой - обеспечивают доступ к полезным функциям из любого места без необходимости настройки VPN-соединения. Среди поставщиков данные системы также популярны, поскольку они на порядок упрощают контроль версий и техническую поддержку (все клиенты, по сути, пользуются одной, наиболее актуальной инсталляцией в «облаке»), а также обеспечивают повторяемые платежи от одних и тех же клиентов, поскольку Интернет-услуги всегда продаются по подписке.

Системы анализа неструктурированной речевой информации (далее - система АНРИ) в значительной степени облегчают задачу аналитических отделов компаний по обработке больших массивов речевых данных, в десятки, а иногда и в сотни раз повышая скорость и качество анализа (подробнее о данных системах см. в [1]). В связи с тем, что популярность данных систем растет, а также с тем, что зарубежные аналоги стоят очень дорого, представляется, что

реализация отечественной версии системы АНРИ в виде WEB-сервиса несет большую потребительскую ценность. В настоящей статье рассматривается принципиальная схема развертывания системы АНРИ в сети Интернет на примере системы поиска ключевых слов в аудиозаписях - коммерческого продукта ANALYZE производства ООО «Спич Драйв».

Состав и архитектура системы

Требования к составу системы определяются исходя из общих требований к любым системам, предназначенным для работы в сети Интернет, а именно: доступность в режиме 24 на 7, безотказность, легкость масштабирования, высокая скорость отклика и надежная система тарификации. Состав системы указан в таблице 1.

Таблица 1 - Компонентный состав Интернет-сервиса системы АНРИ

№	Наименование компонента
1	HTTP-balancer
2	WEB-сервер
3	WEB-интерфейс
4	ПО поиска ключевых слов
5	Сервер нотификаций
6	БД биллинга

Отдельно отметим, что для работы сервиса по защищенному протоколу необходимо приобретение SSL-сертификата и установка его на машину, на которой установлен HTTP-balancer.

Далее полноценный интерфейс для работы с сервисом может реализовываться двумя путями:

а. Сторонней компанией путем поддержки функций https-клиента в своем программном продукте и построения GUI в рамках своего программного продукта

b. Нами путем развития функционала WEB-сервера: генерации html-страниц, позволяющих через Интернет-браузер посмотреть, отредактировать, сохранить на локальную машину результаты поиска.

На рисунке 1 представлена архитектура сервиса. Белым цветом обозначены компоненты сервиса, темно-серым – компоненты, разрабатываемые партнером, светло-серым – вспомогательные инструменты для тестирования сервиса. Пунктиром обозначены резервные компоненты.

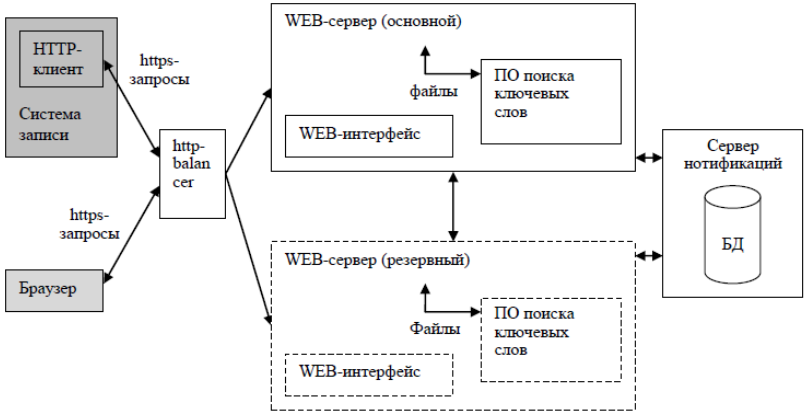


Рисунок 1. Архитектура Интернет-сервиса системы анализа неструктурированной речевой информации

Функциональность компонентов системы

HTTP-balancer

Предназначен для резервирования WEB-серверов и обеспечения тем самым бесперебойной работы сервиса по обработке поисковых запросов клиентов. HTTP-balancer разработан с использованием open-source решения, описанного в [2]. Устанавливается на отдельную машину и выполняет следующие основные функции:

1) Переключение на резервный WEB-Сервер в случае сбоя основного
2) Механизм синхронизации, обеспечивающий безостановочную работу Системы в момент, когда основной WEB-сервер восстановлен и готов воз-

вращать HTTP-клиенту результаты не до конца обработанных запросов.

WEB-сервер

Предназначен для приема от HTTP-клиента (по защищенному соединению) запросов на поиск ключевых слов в звуковых файлах, для передачи запросов к ПО поиска ключевых слов и для возврата HTTP-клиенту результатов обработки его запросов. Так же, как и HTPP-Balancer, WEB-сервер разрабатывается с использованием open-source решений. Существует несколько готовых HTTP-серверов с поддержкой SSL, авторы использовали сервер, описанный в [3]. При выборе севера учитывалась необходимость полной поддержки ОС Windows, т.к. WEB-сервер располагается на той же [виртуальной] машине, что и ПО поиска ключевых слов, которое на текущий момент не поддерживает Linux и иных ОС. WEB-сервер обеспечивает выполнение следующих функций:

1) Взаимодействие с HTTP-клиентом, а именно: создание запроса на обработку данных, передача списка слов и аудиоданных для поиска, привязка данных к уникальному ID запроса, запуск процесса анализа, проверка статуса обработки запроса, остановка и удаление запроса.

2) Взаимодействие с ПО поиска ключевых слов, а именно: запись звуковых файлов и «управляющих» текстовых файлов[1] в папку, мониторинг которой осуществляет ПО поиска ключевых слов. Сами звуковые файлы пишутся под произвольными именами, а запросы формируются отдельным файлом с указанием пар: ключевое слово – имя звукового файла. Такой вариант реализации позволяет произвольно комбинировать ключевые слова и звуковые файлы

[1] *Под управляющими текстовыми файлами подразумеваются xml-файлы, содержащие в себе информацию о запросе. Парсинг данных xml-файлов осуществляется ПО поиска ключевых слов автоматически.*

(поиск 1 слова в нескольких файлах, поиск нескольких слов в 1 файле) без их переименования и копирования.

3) Поддержка защищенного SSL-соединения

4) Поддержка механизма хранения промежуточных результатов обработки запроса с целью возобновления обработки запросов после сбоя

5) Резервирование данных (Резервирование HDD)

6) Скрипт репликации (механизм подгрузки файлов-запросов с резервного сервера при старте основного). На момент репликации оба WEB-сервера автоматически останавливаются.

WEB-интерфейс

Предназначен для тестирования и отладки сервиса. Доступ к WEB-интерфейсу осуществляется через любой стандартный Интернет-браузер. WEB-интерфейс включает в свой состав следующие страницы:

1) Страница идентификации пользователя (логин/пароль);

2) Страница со списком текущих запросов пользователя и выбора действия: создание нового запроса; просмотр статуса одного из текущих запросов; останов одного из текущих запросов;

3) Страницы/формы для загрузки на сервер и удаления списка ключевых слов и звуковых файлов;

4) Страница создания запроса: выбор ключевых слов и звуковых файлов, ранее загруженных клиентом, для данного запроса;

5) Страницы/формы подтверждения запуска/останова/удаления запроса;

6) Страница статуса запроса (текущий процент; остановлен или запущен; ссылка на результат, если поиск завершён; кнопки останова и удаления запроса).

ПО поиска ключевых слов

Предназначено для поиска ключевых слов в звуко-

вых файлах и создания индексных файлов. Используется наиболее актуальная версия ПО ANALYZE разработки ООО «Спич Драйв». Данный компонент обеспечивает выполнение следующих функций:

1) Постоянный мониторинг указанной в конфигурационном файле папки на предмет новых управляющих файлов-запросов от WEB-Сервера
2) Парсинг текстового файла-запроса и поиск указанных ключевых слов в соответствующем звуковом файле
3) Сохранение индексных файлов
4) Механизм сохранения промежуточных результатов, чтобы начинать работу с момента сбоя, а не с начала (анализируется неполный индексный файл)

Сервер нотификаций

Предназначен в первую очередь для оперативной нотификации администратора системы в случае сбоев WEB-сервера. Также может использоваться для мониторинга нагрузки на WEB-сервер с точки зрения количества подключений, количества передаваемых данных, скорость обработки запросов и т.п. Сервер нотификации разработан с использованием open-source решения, описанного в [4], и обеспечивает выполнение следующих базовых действий:

1) E-mail/sms нотификация при сбоях WEB-Сервера
2) Мониторинг нагрузки на WEB-сервер

БД Биллинга

Предназначена для хранения информации об ID пользователей Системы и информации о текущей квоте пользователя (сколько осталось оплаченных часов). Реализуется на базе PostgreSQL и обеспечивает:

1) хранение информации об ID клиентов и их Квотах
2) резервирование БД
3) резервирование HDD

Заключение

В настоящей статье была рассмотрена принципиальная схема развертывания системы анализа неструктурированной речевой информации в сети Интернет. Главная особенность реализации — использование большого числа open-source решений, значительно повысивших скорость развертывания и снижающих риски, связанные с самостоятельной разработкой сложных функций, требуемых для обеспечения качественной работы Интернет-сервиса. В ближайшее время планируется перевести на сетевой протокол связку между WEB-сервером и ПО поиска ключевых слов (вместо текстовых файлов-запросов), ввести масштабирование и модуль распределения нагрузки для ПО Поиска ключевых слов, а также обеспечить более тонкую схему резервирования WEB-серверов за счет использования скриптов репликации.

Список литературы

1. Смирнов В.А., Гусев М.Н., Фархадов М.П. Функция лингвистического процессора в системе автоматического анализа неструктурированной речевой информации // Автоматизация и современные технологии. 2013 (в печати).
2. http://nginx.org/ru/
3. http://jetty.codehaus.org/jetty/
4. http://www.zabbix.com/ru/features.php

УДК 529.12.055
М.В. Гаазе, зам. дир.
Н.И. Лычагин, д-р, техн., наук, зам. ген. дир.
А.А. Малаховский, нач. отдела
ОАО «Информационные телекоммуникационные технологии» (Интелтех),
г. Санкт-Петербург, Россия
nil@ntc1.inteltech.ru

ЛОГИЧЕСКАЯ ОРГАНИЗАЦИЯ МУЛЬТИСЕРВИСНЫХ ОБЪЕКТОВЫХ ВЕДОМСТВЕННЫХ СЕТЕЙ СВЯЗИ

Высокие темпы развития информационных технологий создают объективные предпосылки для предоставления современных телекоммуникационных услуг в автоматизированных системах управления, создаваемых в интересах различных ведомств. Основным местом предоставления услуг связи являются объектовые (офисные) сети, к которым непосредственно подключается терминальное оборудование должностных лиц. В зависимости от особенностей ведомства к информационной безопасности предоставления услуг связи в объектовых сетях и организации удаленного взаимодействия объектовых сетей предъявляются различные требования. В данной работе реализация требований к информационной безопасности не рассматривается.

Современные услуги связи включают в себя широкий спектр мультисервисных услуг от громкоговорящей связи до потокового телевидения. Типичной интегрированной услугой является видеоконференцсвязь (ВКС), объединяющая передачу голоса, видео и данных. В дальнейшем будем рассматривать реализацию мультисервисных услуг на примере ВКС.

Ситуационные центры верхнего уровня иерархии в АСУ многих ведомств к настоящему времени оснащены

зарубежным оборудованием ВКС производства ведущих мировых фирм, таких как Tandberg и Polycom.

Высокая стоимость зарубежного оборудования ВКС, отсутствие открытого программного кода, разнотипность установленных средств существенно затрудняют оснащение средних и нижних уровней АСУ требуемыми терминальными и серверными средствами. При этом эффективность функционирования АСУ снижается, так как отсутствуют необходимые средства для передачи оперативной мультимедийной информации между уровнями иерархии системы. Комплексное оснащение различных уровней АСУ оборудованием, предоставляющим современные интегрированные телекоммуникационные услуги, целесообразно проводить на основе отечественных разработок, ориентированных на строгое соблюдение перспективных международных стандартов и рекомендаций, имеющих четкую документированную системную идеологию, распространяющуюся на все уровни иерархии АСУ, и обладающих возможностями адаптации к требованиям различных ведомств по обеспечению информационной безопасности, включая работу в различных операционных средах. Остановимся на основных проблемных вопросах логической организации рассматриваемых сетей.

Выбор протокольного стека и нормативная база

Сегодня основными концепциями развития телекоммуникационных сетей являются NGN (Next-Generation Network) и IMS (IP Multimedia Sub-system), в рамках которых реализация телекоммуникационных услуг базируется на единой технологии управления сеансами – протоколе SIP (Session Initiation Protocol).

Большинство создаваемых в последнее время систем телефонии и видеотелефонии в корпоративных и ведомственных сетях связи придерживается основного направления развития и является SIP-ориентированными. Однако в качестве систем аудио и видеоконференцсвязи до сих пор используют системы на базе протоколов H.323.

Типовой режим работы таких конференций – централизованный. Потоки данных, как управления, так и медиа (аудио, видео) проходят непосредственно между каждым терминалом и сервером управления конференцией (MCU). Наиболее высокая нагрузка в этом варианте ложится на медиасервер, обеспечивающий обработку медиапотоков. По этой причине возможности по организации сессии конференции (по количеству участников и качеству передаваемых данных) ограничены возможностями MCU. Его наличие снижает также надежность таких систем.

Современным сигнализационным протоколом, обеспечивающим управление вызовами, является протокол SIP, который пришел на замену H.323.

Стек протоколов SIP обеспечивает органичную и естественную поддержку возможности организации распределенных конференций с децентрализацией передачи и обработки медиаданных. Такой подход позволяет избавиться от дорогого и сложного с технической точки зрения устройства (медиасервер) и резко повысить надежность системы в целом за счет устранения единой точки отказа и увеличения топологической надежности системы. Однако при этом повышаются требования по производительности терминального оборудования и пропускной способности каналов связи. Оба фактора в настоящее время перестают быть препятствием к динамичному развитию распределенных технологий в связи с постоянным увеличением возможностей базовых средств вычислительной техники и увеличением пропускной способности каналов.

Протокол SIP обеспечивает также полнофункциональную поддержку каскадных конференций. Причем каскадирование может осуществляться как в части обработки медиаданных, так и в части управления сессиями конференций. Возможны любые сочетания этих режимов в зависимости от характеристик сети связи и территориально-организационных аспектов создаваемой системы. Совместное использование технологий децентрализации и каска-

дирования обеспечивает динамическое управление медиапотоками в сессии конференцсвязи, позволяет в наибольшей степени оптимизировать передачу медиаданных по сетям с любой топологией и эффективно использовать ресурсы неоднородной в части пропускных способностей каналов сети. При этом медиасервер может выполнять различные группы функций, начиная от микширования получаемых от различных источников медиаданных и заканчивая процедурами распределения медиапотоков в сети без транскодирования и промежуточной обработки.

Применение таких технологий обеспечивает возможность организовывать так называемые «иерархические» конференции, в которых каждая сессия может состоять из одной основной и нескольких дочерних так называемых рубрик. Рубрики позволяют организовать дополнительную передачу медиаданных параллельно с данными основной конференции, причем данные участников рубрики недоступны участникам основной сессии и участникам других рубрик. Глубина вложенности таких рубрик может быть практически неограниченной.

Рубрики могут применяться для организации вспомогательных консультативных конференций или для получения дополнительной информации в реальном масштабе времени от информационных систем различного назначения.

Исходя из представленных выше преимуществ, в качестве протокола сигнализации для программной реализации услуги видеоконференцсвязи был выбран протокол SIP. Для организации и управления сессиями конференцсвязи использован ряд рекомендаций организации IETF, которые можно разбить на следующие области:

– архитектура конференц-связи: RFC 4245, RFC 4353;

– управление вызовом третьей стороной: RFC 3725, RFC 4579;

– управление выступлениями: RFC 4376, RFC 4582,

RFC 4583;

– система уведомлений: RFC 3265, RFC 4575.

Выбор протоколов для передачи по сети аудио- и видеоинформации является достаточно тривиальным, протоколы RTP и RTCP являются общепринятыми международными стандартами для решения этой задачи в IP-сетях с любыми видами сигнализации. Форматы и структуры сообщений протоколов RTP и RTCP, а также порядок их следования и обработки представлены в рекомендации RFC 3550. Использование этих протоколов для передачи данных при различных способах кодирования аудио- и видеоинформации регламентируются рядом рекомендаций, таких как RFC 3551, RFC 4587, RFC 2190, RFC 3984, RFC 3016, RFC 3640 и т.д.

В рамках международных стандартов услуги передачи мгновенных сообщений и индикации статусов пользователей в сети (услуга "Присутсвие") с использованием протокола SIP увязаны в единую систему SIMPLE (SIP for Instant Messaging and Presence Leveraging Extensions), обеспечивающую наибольшие функциональные возможности по сравнению с системами, базирующимися на других протоколах (H.323, XMPP и т.д.).

Кроме стандартной статусной информации PIDF (Presence Information Data Format) с использованием SIP может передаваться дополнительная информация, включающая расширения RPID (Rich Presence Extensions to the Presence Information Data Format), CIPID (Contact Information for the Presence Information Data Format), GEOPRIV. RPID позволяет индицировать текущее состояние и возможности оконечных устройств, а также обеспечивает привязку статуса к стандартным адресным книгам, календарям и плановикам. CIPID описывает расширения PIDF в части представления абонента, позволяющего воспроизводить графическую и аудиоинформацию, связанную с индицируемым пользователем. GEOPRIV позволяет описывать географическое местоположение с использованием языка

GML (Geography Markup Language).

Для организации дополнительных услуг по передаче данных существует множество различных подходов и стандартных решений, однако представляется целесообразным использование решений, основанных на выбранном стеке протокола сигнализации SIP. Доставка текстовых сообщений с использованием протокола SIP может быть выполнена в двух режимах: пейджинговом и сессионном. Пейджинговый режим передачи мгновенных сообщений подразумевает доставку данных в отложенном режиме с минимальными гарантиями. Данный режим описан в документе RFC 3428. Сессионный режим обеспечивает гарантированную доставку сообщений во времени, близком к реальному, и реализуется при помощи протокола MSRP, специфицированного в рекомендации RFC 4975. Для передачи файлов целесообразно использовать только сессионный режим, который реализуется с использованием протокола MSRP.

Таким образом, анализ нормативной базы показывает, что широкий спектр телекоммуникационных услуг, связанных с предоставлением возможности обмена мультимедийной информацией, включая аудио- и видеоданные, может быть реализован на базе единого сигнализационного протокола в соответствии с международными стандартами и рекомендациями (см. рисунок 1).

Функциональная модель предметной области

Целью создания функциональной архитектуры системы является декомпозиция ее общей функциональности на ряд функциональных модулей, каждый из которых реализует группу тесно взаимосвязанных функций. Такое разбиение можно делать различными способами, в зависимости от выбранных критериев. Предлагаемая функциональная архитектура позволяет минимизировать и достаточно четко специфицировать связи между отдельными функциональными модулями системы, реализующей предостав-

ление телекоммуникационных услуг на базе протокола сигнализации SIP. Это, в свою очередь, обеспечивает возможность независимой проработки различных аспектов реализации каждого из функциональных модулей и возможность в будущем замены одного из них без существенного влияния на остальные.

Общую функциональность системы представляется целесообразным разделить на функциональные модули и функциональные плоскости. Такое разделение позволяет представлять процессы, проходящие в рамках системы, в виде взаимодействия функциональных модулей. Причем взаимодействие модулей группируется по функциональным областям на функциональных плоскостях.

При разработке функциональной архитектуры были выделены функциональные модули и плоскости, представленные на рисунке 2. Рисунок 2 также отражает распределение функциональных модулей между функциональными плоскостями. Наиболее функционально насыщенными являются модули терминал и пограничный контроллер сессий, которые участвуют в работе всех представленных функциональных плоскостей. Среди функциональных плоскостей наиболее насыщенной является плоскость сигнализации, которая объединяет функции всех без исключения функциональных модулей.

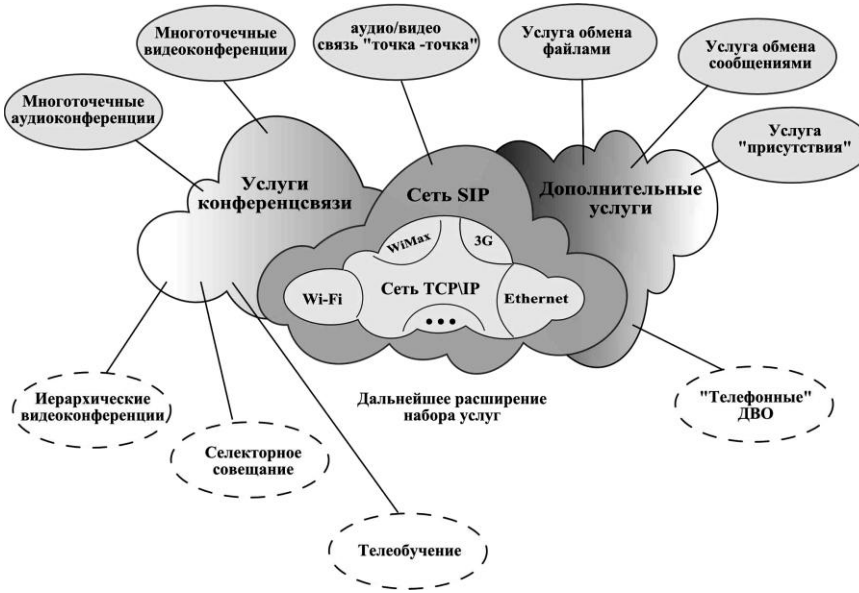


Рисунок 1 – Возможности платформы предоставления мультисервисных услуг связи

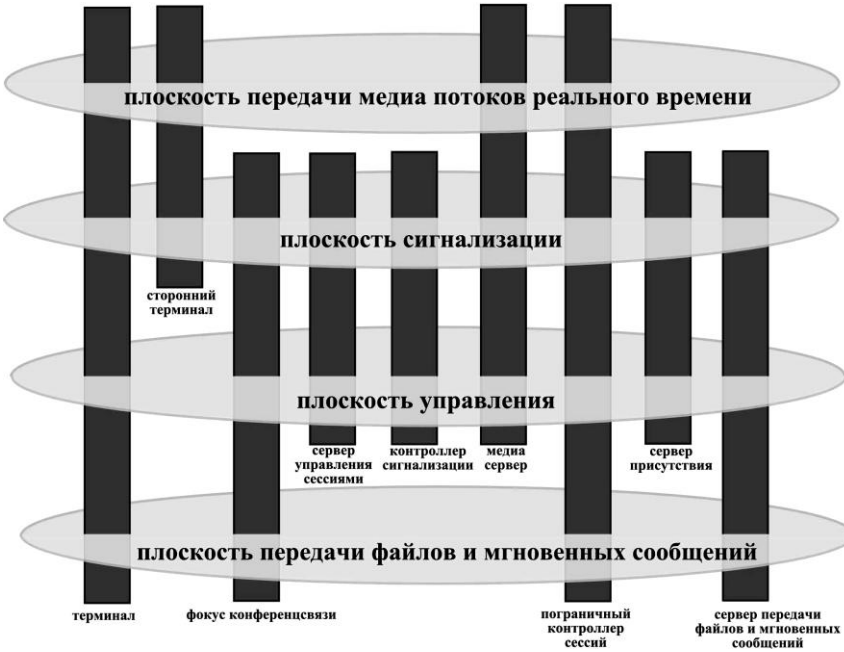


Рисунок 2 – Распределение функциональных модулей по функциональным плоскостям

Рассмотрим более подробно функциональность каждого из перечисленных выше модулей и связи между этими модулями в рамках функциональных плоскостей.

Плоскость сигнализации описывает процессы взаимодействия функциональных модулей в части управления вызовами. Сеть сигнализации должна обеспечивать управление предоставлением базовой услуги аудио- и видеосвязи, но также реализует и управление дополнительными услугами.

Ведомственные сети связи целесообразно разделять на домены, каждый из которых может иметь независимое управление и администрирование. В этом случае в рамках плоскости сигнализации необходимо выделить два вида взаимодействия функциональных модулей: внутридоменное и междоменное. Внутридоменное взаимодействие обеспечивается контроллером сигнализации. Междоменное

взаимодействие осуществляется при помощи пограничного контроллера сессий.

За управление вызовами в рамках сессии конференц-связи отвечают терминал, сервер управления сессиями конференцсвязи и фокус конференц-связи. Терминал является оконечным узлом сети сигнализации и выполняет функции подключения и отключения от сессии наряду с обработкой управляющей информации о медиапотоках участника в сессии, а также обработкой дополнительной информации о сессии, включая список участников.

Сервер управления сессиями конференц-связи отвечает за запуск и завершение работы фокусов конференц-связи. Последние отвечают за управление медиапотоками в сессиях. Возможны два режима по управлению фокусом конференц-связи:

– централизованный;

– децентрализованный.

В централизованном режиме за запуск и завершение сессий отвечает сервер управления сессиями, который производит запуск и завершение фокуса конференц-связи. В децентрализованном режиме фокус конференц-связи является частью терминала и управляется пользователем.

Наличие дополнительных услуг в сети обеспечивается функциональными модулями сервера присутствия и сервера передачи файлов и мгновенных сообщений.

Плоскость передачи медиапотоков реального времени описывает процессы взаимодействия функциональных модулей в части приема, передачи и обработки аудио- и видеопотоков в рамках сессии связи. Основными функциональными модулями, обеспечивающими работу данной плоскости, являются терминал, медиасервер и пограничный контроллер сессий.

Терминал обеспечивает предоставление пользователю услуги по формированию и воспроизведению потоков от участника сессии. Медиасервер выполняет функции мультиплексирования потоков, подключения сторонних

терминалов, а также может обеспечивать создание и формирование дополнительных медиапотоков в сессии конференц-связи. Пограничный контроллер сессий отвечает за междоменное взаимодействие остальных модулей.

Передача медиапотоков реального времени в рамках сессии конференц-связи может проходить в следующих режимах:

– централизованный;

– децентрализованный;

– смешанный.

В централизованном режиме все терминалы передают потоки реального времени на медиасервер, который обеспечивает их микширование и распределение мультиплексированного потока ко всем терминалам участников сессии конференц-связи.

В децентрализованном режиме медиасервер не используется. Терминалы передают потоки реального времени непосредственно между собой по схеме «каждый с каждым».

Смешанный режим позволяет передавать часть потоков децентрализовано, а часть – централизовано через медиасервер.

Плоскость передачи файлов и мгновенных сообщений описывает процессы взаимодействия функциональных модулей в части приема, передачи и обработки данных. Основными функциональными модулями, обеспечивающими работу данной плоскости, являются терминал, сервер передачи файлов и мгновенных сообщений, фокус конференц-связи и пограничный контроллер сессий.

Терминал в плоскости передачи файлов и мгновенных сообщений отвечает за формирование интерфейса пользователя по приему и передаче файлов и мгновенных сообщений другим пользователям. Сервер передачи файлов и мгновенных сообщений обеспечивает режим отложенной передачи данных. Фокус конференц-связи в данной плоскости отвечает за передачу файлов и мгновенных со-

общений группе участников сессии. Пограничный контроллер сессий обеспечивает междоменное взаимодействие в плоскости передачи файлов и мгновенных сообщений.

Анализ представленной функциональной модели показывает, что подавляющее большинство связей между функциональными модулями основано на использовании сетевых протоколов (SIP, RTP, MSRP и т.д.). Это означает, что физические реализации этих модулей могут быть при необходимости распределены между различными компьютерами сети. Исключение составляет только фокус конференц-связи, который должен быть встроен в сервер управления сессиями или в терминал.

Предлагаемая функциональная модель ориентирована в первую очередь на создание перспективных распределенных телекоммуникационных служб. При этом она допускает различные степени децентрализации, которые условно можно разделить на четыре уровня:

– полностью распределенная система, серверное оборудование любой функциональ-ности отсутствует в ее составе, терминалы могут самостоятельно организовывать многоточечные сессии конференц-связи, при этом один их них управляет сессией, медиадан-ные передаются по принципу «от каждого к каждому»;

– централизация по обработке сигнализационных сообщений, в системе появляется сервер сигнализации (SIP-прокси), что обеспечивает поддержку функций авторизации, аутентификации, мобильности пользователей, а также возможность работы с символическими адресами;

– централизация по управлению сессиями конференц-связи, система наращивается сервером конференции, что обеспечивает возможность ведения централизованных адресных книг и сценариев, позволяет запускать сессии конференц-связи в автоматическом режиме по расписанию, позволяет централизованно осуществлять сбор статистики (предбиллинг) по сессиям конференц-связи;

– централизация по обработке медиаданных, система дополнительно наращивается медиасервером, что позволяет понизить требования к производительности терминального оборудования и пропускной способности каналов связи, а также в некоторых случаях более гибко учитывать топологические особенности сети связи.

Предлагаемая логическая организация была апробирована при разработке системных проектов мультисервисных объектовых сетей нескольких ведомств и реализована в виде модульного аппаратно-программного комплекса, который внедряется в ряд ведомств [1, 2].

Список литературы

1. Гаазе М.В., Лычагин Н.И., Малаховский А.А. Методология создания и реализации интегрированных телекоммуникационных служб для ведомственных сетей связи // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. 2012. № 6 (162). - С. 52-58.

2. Ершов В.Н., Катанович А.А., Лычагин Н.И. Современный подход к построению внутрикораблеьной системы связи // Судостроение. 2013. № 2. - С. 22-25.

УДК 556.54.18:51; 681.3+ 004.4 + 004.738
И.М. Гуревич, канд. техн. наук
Институт проблем информатики РАН, ООО «ГЕТНЕТ Консалтинг»,
г. Москва, Россия, iggurevich@gmail.com
Л.К. Левит-Гуревич, канд. техн. наук
Институт водных проблем РАН, г. Москва, Россия
Lev-Gur@yandex.ru

ИСПОЛЬЗОВАНИЕ МОДЕЛЕЙ СЕТЕВЫХ СИСТЕМ ДЛЯ ОПИСАНИЯ И ОЦЕНОК ГИДРОГРАФИЧЕСКИХ СЕТЕЙ

Введение

Водное хозяйство является одной из важных отраслей экономики любой страны, в особенности, России, поскольку Россия занимает второе место в мире по объемам водных ресурсов (после Бразилии) и первое место по числу и разнообразию водных объектов. Поверхностные воды рек и озер являются источником водоснабжения, в частности, питьевого водоснабжения (здесь большую роль играют подземные воды), также деятельности рыбного хозяйства, гидроэнергетики, ирригации, водного транспорта. Поэтому проблем и сформулированных задач, связанных с водными ресурсами, очень много, и известны они многие годы (некоторые и века), решения их развивались в соответствии с развитием математики и вычислительной техники в последние десятилетия.

С другой стороны, развитие сетей связи, в частности, телефонии, способствовало появлению таких математических дисциплин, как теория очередей, и развитию таких областей знания, как сетевые модели, потоковые модели, Марковские цепи, … Появляется мысль использовать аппарат этих дисциплин для решения задач водного хозяйства, поскольку водохозяйственные системы (в принципе, все водное хозяйство) как и связь, является являются сете-

выми системами. Водохозяйственные системы аналогичны сетям связи, с теми же балансовыми законами, но есть и существенное различие, назовем гидравлику, потери воды на испарение, фильтрация и пр.

Большой интерес представляют задачи ресурсораспределения. В вершину *A* поступает объем некоторого ресурса, – его надо распределить не по одному маршруту дуг, а по всем дугам графа *G* до вершины *B*, сообразуясь с разными оценками пропуска ресурса по дугам и оптимальным значением целевой функции из этих оценок. Задачи ресурсного распределения на графе зачастую отличаются линейностью условий по дугам и балансовыми условиям в вершинах. Такие задачи решаются с использованием методов оптимизации.

В настоящем докладе делается попытка показать, что в водных проблемах и задачах могут быть использованы модели сетевых систем, в частности, решение задач оптимизации может упроститься при использовании для описания каскада водохранилищ многоуровневой статической и динамической моделей сетевых систем. Нельзя сказать, что некоторые эти модели не использовались ранее в водном хозяйстве, но интересен общий взгляд, общие положения (принципы) и общие, наиболее эффективные приемы.

Необходимые понятия и определения

В связи с тем, что авторы предлагают новое применение сетевых моделей, целесообразно привести определения понятий водного хозяйства [1-11].

Во́дный объе́кт – сосредоточение вод на поверхности суши в формах её рельефа либо в недрах, имеющее границы, объём и черты гидрологического режима. К водным объектам относятся водоемы и водотоки. Водными объектами являются моря, океаны, реки, озёра, болота, водохранилища, а также воды каналов [1].

Гидрографи́ческая сеть – совокупность рек и других постоянно и временно действующих водотоков, а так-

же озёр, болот и водохранилищна какой-либо территории. Когда рассматривается только система водотоков, применяется термин «речная сеть». Однако нередко понятия *гидрографическая сеть* и *речная сеть* отождествляются [2].

Прито́к – водоток, впадающий в более крупный водоток. Различают притоки разных порядков в зависимости от того, впадают ли они непосредственно в главную реку или в её притоки. Притоками первого порядка называются реки, непосредственно впадающие в главную реку, второго порядка – притоки притоков первого порядка и т.д. Крупные речные системы включают до 20 порядков притоков. Существуют противоположные подсчёты порядка притоков, - от мелких к крупным. Также **притоком** называют расход воды, приносимый водотоками в озеро, водохранилище и другие водные объекты [3].

Рука́в – сформировавшееся отдельное русло реки со всеми свойственными речному руслу особенностями морфологического строения. Рукав обычно вытекает из основной реки и впадает в неё же ниже по течению. В низовьях крупных рек, как правило, из множества рукавов, впадающих затем в море, образуется **дельта** [4].

Бассе́йн водоёма (водосборная площадь) – территория земной поверхности, с которой все поверхностные и грунтовые воды стекаются в данный водоём, включая различные его притоки. Граница между бассейнами отдельных водоёмов проходит по **водоразделам** - условная линия на карте земной поверхности, разделяющая водосборные бассейны рек, озёр, морей. Бассейны бывают сточные и бессточные. Бессточными называются области внутриматерикового стока, лишённого связи через речные бассейны с океаном [5].

Уровень воды – высотная отметка воды Z м у гидротехнического сооружения, в створе реки [6].

Расход воды (в водотоке) – объём воды $\boldsymbol{Q}$, протекающей через поперечное сечение водотока за единицу времени. Измеряется в расходных единицах ($м^3/с$). В об-

щем случае методология измерения расхода воды в реках и трубопроводах основана на упрощённой форме уравнения непрерывности, для несжимаемых жидкостей: $Q=Fv$, где F м^2 - площадь поперечного сечения водотока (трубы или части русла реки, заполненного водой), $\bar{v}$ - средняя скорость потока [м/с] [7].

Гидрологи́ческий режи́м – изменения состояния водного объекта, описываемее уровнями и расходами во времени (и в пространстве данного объекта). Естественный гидрологический режим стохастический, деятельность человека, скажем, регулирование стока реки водохранилищем, вносит закономерности в гидрологический режим реки и водоема [8].

Гидрология – научная дисциплина, рассматривающая стохастические законы движения воды движение воды в природе: расходы протока в водоем, речной сток и др. Изучает вероятностные распределения уровней или расходов по времени [9].

Гидравлика – научная дисциплина, рассматривающая движение воды в естественных руслах с поймой или в трубопроводах. Гидравлика рассматривает установившееся и неустановившееся движения воды, первое описывается в естественном русле уравнением Шези нелинейной связи уровня воды и расхода, второе – системой дифференциальных уравнений Сен-Венана гиперболического типа с разными методами интегрирования [10].

Расчетная обеспеченность, - величина вероятности q^* на кривой вероятности $Q(q)=\int_{-\infty}^{q} g(\eta)d\eta$, где $g(\eta)$ – плотность вероятности некоторой интересующей практику величины. Расчетная обеспеченность определяет то значение величины Q, которое используется при проектировании водохозяйственной системы, сооружения или мероприятия. Например, высота дамб обвалований, защищающих населенные площади или сельскохозяйственные угодья от наводнений, зависит от максимальных расходов, проходящих в период половодья, и максимальных уровней воды; выбор

высоты дамб представляет собой технико-экономи-ческую проблему, в которой выбор расчетной обеспеченности имеет ключевое значение [11].

Все проектные задачи водного хозяйства – это оптимизационные задачи, использующие для своего решения методы оптимизации. В частности, большой интерес представляют задачи ресурсораспределения. В вершину *A* поступает объем некоторого ресурса, – его надо распределить не по одному маршруту дуг, а по всем дугам графа *G* до вершины *B*, сообразуясь с разными оценками пропуска ресурса по дугам и оптимальным значением целевой функции из этих оценок. Задачи ресурсного распределения на графе зачастую отличаются линейностью условий по дугам и балансовыми условиям в вершинах. Такие задачи решаются с использованием идеологии потоковых моделей [12], методами линейного программирования (ЛП), нелинейного программирования (НЛП), динамического программирования (ДП) [13-15]. Решение задач оптимизации может существенно упроститься при использовании для описания гидрографической сети многоуровневой статической и динамической моделей сетевых систем. Опишем многоуровневые статическую и динамическую модели сетевых систем [16-18].

Многоуровневая статическая модель сетевой системы

Предлагаемая трехуровневая модель сетевой системы модели сетевого уровня, модели уровня узла, модели уровня ребер. Модели сетевого уровня отображают процессы передачи информации по сети связи в целом, описывают вероятностно-временные и прочие характеристики сети. Модели уровня узла отображают процедуры выбора маршрутов и управления потоками, процедуры распределения потоков между ребрами. Описывают характеристики узлов сети. Модели уровня ребер отображают процедуры обслуживания потоков в ребрах. Описывают характеристики ребер.

Законы сохранения в сетях находят свое выражение, прежде всего в системе потоковых уравнений. Системы потоковых уравнений составляются относительно интенсивности потоков, поступающих на вход элементов, фрагментов сети и отражают тот факт, что потоки в процессе передачи по сети связи не могут самопроизвольно появляться и исчезать, что поток в соответствии с реализуемыми сетью сетевыми алгоритмами из данного узла (ребра) будет доведен до адресата или передан в другой узел (ребро), или будет «потеряна» вследствие отказов в поступлении в ребра, истечение допустимого времени пребывания в сети и т.п. Потоковые уравнения могут составляться как для суммарных потоков, обладающих определенными свойствами или отличительными признаками, например, для потоков заданных приоритетов, потоков, идущих в определенные адреса, ребра. Потоковые уравнения могут взаимосвязывать потоки в смежных и в произвольных элементах сети. Системы потоковых уравнений составляются на всех уровнях рассматриваемой модели сети. Системы потоковых уравнений сетевого уровня $\lambda^t{}_j = \Sigma_i P_{i\,j} \lambda_i + V^t{}_j$, где i, j, t = 1,…, n. Данная система уравнений говорит о том, что интенсивность потока, идущего по адресу t и поступающего на узел j, равна сумме интенсивности потоков, поступающих в данный узел j из других узлов i – $P^t{}_{i\,j} \cdot \lambda_i$. и внешнего потока, поступающего в данный узел j – V^t_j. В состав рассматриваемых узлов включаются узлы сети, которым соответствуют состояния конечных цепей Маркова [19-20], а также дополнительные (фиктивные) состояния, в которые в модели направляются «потерянные» сообщения. P^t_{ij} – элементы матриц переходных вероятностей процессов P_t. Для однопродуктовой сети индекс t можно опустить и получим системы уравнений: $\lambda_j = \sum P_{i\,j} \lambda_i + V_j$. Вид систем потоковых уравнений определяется видом матриц переходных вероятностей P_t, сетевыми алгоритмами и правилами обслуживания очередей. Если, например, в сети используются сетевые алгоритмы передачи

информации по фиксированным маршрутам или случайные процедуры выбора направлений передачи, а ребра бесконечны, то P^t_{ij} = const и системы потоковых уравнений являются линейными системами уравнений. Если в сети используются простые или адаптивные алгоритмы, то в рамках рассматриваемой модели, системы потоковых уравнений являются нелинейными алгебраическими уравнениями. Размерность систем потоковых уравнений сетевого уровня определяется числом узлов сети, числом адресатов и пропорциональна n t (n – для однопродуктовой сети).

Модели сетевого уровня отображают процессы передачи информации по сети в целом. Сетевой уровень представляется совокупностью P конечных полумарковских процессов (КПП) [19-20]. Каждому узлу адресату t ($I \le t \le n$) ставится в соответствии один и только конечный полумарковский процесс P_t: $P = U_t P_t$. Состояния КПП P_t в общем случае отождествляются с узлами сети. Как правило, все процессы P_t определенны на одних и тех же состояниях. Каждому узлу может ставиться в соответствие несколько состояний КПП. При необходимости, состояниями КПП могут объявляться определенные совокупности элементов сети, ребра, события. Могут также вводиться дополнительные или фиктивные состояния (в которые, например, будут направляться потерянные по разным причинам потоки). Определение множеств(а) состояний процессов P_t производится Исследователем или Конструктором в соответствии со стоящей перед ним задачей. Элементы матриц переходных вероятностей и значения процессов на состояниях КПП (узлах сети) вычисляются моделями уровня узла. Формулы для вычисления сетевых характеристик приведены ниже. Модели уровня узла отображают процедуры выбора маршрутов и управления потоками, процедуры распределения потоков между ребрами. Модели уровня узла взаимосвязывают значения вероятностей P^t_{ij} передачи потоков между узлами сети (P^t_{ij}

являются элементами переходных вероятностей P_t КПП P_t), значения вероятностей P^{ijt}_{pq} выбора направлений pq (направления pq отождествляются с ребрами pq, времени пребывания потоков в узлах сети с вероятностями a_{mf} направления потоков в ребра mf и временем обслуживания потоков. Кроме того, модели уровня узла определяют зависимости вероятности a_{mf} направления потоков в ребра (направление mf) от вероятностно-временных характеристик обслуживания потоков и сетевых констант æ, влияющих на выбор направ-лений передачи. Модели уровня очередей могут быть заданы в виде формулы, систем ура-внений, решение которых производится в ходе вычисления значений сетевых характеристик, в виде параметров или функциональных зависимостей, полученных в результате имитацион-ного моделирования или статистических измерений.

Приведем ряд основных формул теории конечных цепей Маркова, с использованием которых производится оценка сетей. $N = (I-Q)^{-1}$ - фундаментальная матрица системы, где Q - матрица, получаемая из матрицы переходных вероятностей системы P вычеркиванием строки и столбца соответствующих части системы, отождествленной с поглощающим состоянием; I - единичная матрица; $N = (n_{ij})$, где n_{ij} - число попаданий потока выходящего из части i системы в часть j; $\tau = N\ \theta$, где $\tau = (\tau_i)$ - вектор-столбец средних значений про-извольной аддитивной характеристики системы $\theta=(\theta_i)$. θ_i - значение аддитивной характе-ристики системы на i-й ее части. $D = (d_i) = (2N-I)N \bullet \theta_{sq} - \tau_{sq}$, где (d_i) – вектор-столбец дисперсией аддитивной характеристики τ. A_{sq} матрица, полученная из матрицы A воз-ведением ее элементов в квадрат. $B=(b_{ij})=NR$, где b_{ij} - вероятность попадания потока выходящего из части i в часть системы j (доля потока попавшего в j-ю часть системы); R - вектор-столбец вероятностей перехода из частей системы в часть отождествленную с поглощающим состоянием. $H = (h_{ij}) =$

$(N\text{-}I)N_{dq}^{-1}$, где h_{ij} - вероятность попадания потока, выходящего из части i системы в часть j; A_{dq} - матрица, полученная из матрицы A, заменой всех недиагональных элементов нулями. Объем вычислений по приводимым формулам пропорционален n^3.

Динамическая модель сетевой системы

Реальные сетевые элементы систем связи и управления являются нестационарными объектами. Их нестационарность определяется, в частности, изменениями нагрузки, перемещениями абонентов сети, выходом из строя и восстановлением элементов сети, ее реконфигурацией и т. п. Отсутствие адекватного, по возможности несложного, аппарата исследования нестационарных сетей приводит к тому, что сети рассчитываются на максимальные пиковые нагрузки. Кроме того, при проектировании сетей не проводится анализ их устойчивости, времени и качества переходных процессов. Динамические свойства сети (изменение потоков в узлах) в точке $\lambda_o = (\lambda_{io})$ определяются статическими характеристиками сети, стационарными интенсивностями λ_{io} потоков в узлах сети, вероятностями $p_{ij}(\lambda_{io})$ передачи потока между узлами сети, плотностями времен пребывания потока в ребрах сети $\varphi_{ij}(t, \lambda_{io})$ и изменением интенсивностей входных потоков $v_j(t)$.

Рассмотрим ребро ij. Предположим, что на вход ребра в момент времени t_0 поступает стационарный поток λ°_i, на выходе ребра - стационарный поток λ^{o}_{ij}. Пусть $\Phi_{ij}(\tau)$ - функция распределения времени пребывания потока. $\Phi_{ij}(\tau)$ обладает следующими свойствами: 1) Функция распределения времени пребывания в ребре $\Phi_{ij}(\tau)$ зависит от интенсивности λ°_i потока, поступающих на вход ребра ij $\Phi_{ij}(\tau) = \Phi_{ij}(\tau, \lambda^\circ_i)$. Эта зависимость, как правило, является нелинейной. 2) $\Phi_{ij}(\tau,\lambda^\circ_i) = 0$, при $\tau \le 0$. 3) $\lim \Phi_{ij}(\tau, \lambda^\circ_i) = 1$ при $\tau \to \infty$. Свойства 2, 3 выполняются во всей области Ω_i существования функции распределения $\lambda^\circ_j \subset \Omega_i$. Предполагаем существование плотности распределения времени в

ребре $\varphi_{ij}(\tau, \lambda^{\circ}_i) = \partial\Phi_{ij}(\tau, \lambda^{\circ}_i)/\partial\tau = \Phi'_{ij}(\tau, \lambda^{\circ}_i)$. Рассмотрим ребро под воздействием ступенчатого изменения интенсивности входного потоки. Функция распределения времени обслуживания потока в ребра ij – $\Phi_{ij}(\tau, \lambda^{\circ}_i)$ есть переходная функция ребра $\lambda_{ij}(\tau, \lambda^{\circ}_i) = \Phi_{ij}(\tau, \lambda^{\circ}_i) \cdot \lambda_i(0)$. Рассмотрим изменения интенсивности выходного потока $\lambda_{ij}(t)$ ребра ij.

При $t \le t_0$ $\lambda_{ij}(t) = 0$; при $t > t_0$ $\lambda_{ij}(t) > 0$; при $t \to \infty$ $\lambda_{ij}(t) \to \lambda^{\circ}_i$.

Теперь рассмотрим ребро под воздействием произвольного малого изменения входного потока. Использование $\Phi_{ij}(\tau) = \Phi_{ij}(\tau, \lambda^{\circ}_i)$ в качестве переходной функции позволяет описать изменение выходного потока при произвольном (но малом) изменении входного. Для этого достаточно использовать представление выходного потока в виде интеграла Дюамеля $\lambda_{ij}(t) = d/dt \int_o^t \Phi_{ij}(t-\tau)\, \lambda_i(\tau)\, d\tau = d/dt \int_o^t \Phi_{ij}(\tau)\, \lambda_i(t - \tau) d\tau$. Последнее выражение удобнее представить в следующем виде: $\lambda_{ij}(t) = \Phi_{ij}(0, \lambda^{\circ}_i)\, \lambda_i(t) + \int_o^t \Phi'_{ij}(t-\tau)\, \lambda_{ij}(\tau)\, d\tau$. Ввиду того, что $\Phi_{ij}(0, \lambda^{\circ}_i) = 0$ и $\Phi'_{ij}(\tau) = \varphi_{ij}(\tau)$, получаем выражение $\lambda_{ij}(t) = \int_o^t \varphi_{ij}(t - \tau)\, \lambda_i(\tau) d\tau$.

Учет временной задержки потока в ребрах сети приводит к интегральным уравнениям Вольтерра второго рода $\lambda_{ij}(t, v_j(t)) = \int_o^t \varphi_{ij}(t - \tau, \lambda_{io})\, \lambda_i(\tau, \lambda_{io})\, d\tau$. Динамика сетевой системы в целом описывается системой нелинейных интегральных уравнений Вольтерра второго рода $\lambda_j(t, \lambda_{io}) = \sum_i \lambda_{ij}(t, \lambda_{io}) = \sum_i p_{ij}(\lambda_{io}) \int_o^t \varphi_{ij}(t - \tau, \lambda_{io})\, \lambda_i(\tau, \lambda_{io})\, d\tau + v_j(t)$.

Описанные выше модели сетевых систем авторы предлагают использовать для моделирования гидрографических сетей.

Принципы и пример моделирования гидрографической сети

Введем необходимые обозначения: i – водоем (водохранилище, озеро, ильмень), узел сети. i = 1, 2, … , I. ij –

водотоки (притоки, рукава), смежные дуги.

На узлы сети i поступают внешние потоки воды v_i Из из узла сети i в водотоки ij поступают потоки λ_{ij}.

Для учета расхода, потерь воды необходимо добавить «фиктивные» состояния, на-пример, состояние I+1, в которое направлять всю израсходованную, потерянную воду. Если необходимо различать разные типы расходов, потерь, то необходимо добавить несколько «фиктивных» состояний, например, состояния I+1, I+2, … , I+k, каждое из которых будет характеризовать конкретный тип расходов, потерь.

λ_i – суммарный поток, входящий в i-ый узел. Поток λ_i, в общем случае равен сумме потоков идущим по всем водотокам – смежным дугам ij λ_{ij}, плюс потерянные потоки, плюс внешний поток v_i. $\lambda_i = \sum_j \lambda_{ij} + v_i$. Пусть p_{ij} - доля потока, либо выходящего из водоема i по водотоку ij, либо потерянного очевидно, что $\sum_j p_{ij} = 1$. $\lambda_{ij} = p_{ij}\lambda_i$. Расходы воды на водоеме i связан с забором её на нужды водоснабжения, на потери, а также на пополнение запасов воды в водоеме (хранилище, озере, ильмене).

Тем самым мы фактически определили конечную цепь Маркова [18, 19]. Состояния сети i отождествляются с узлами сети (водоемами), матрица переходных вероятностей $P = \left(p_{ij} \right)$ характеризует связь потоков в узлах и ребрах сети. Матрица переходных вероятностей $P = \left(p_{ij} \right)$ является стохастической матрицей [18, 19] $\sum_j p_{ij} = 1$.

Описываемые далее модели дают возможность оценить, например, потоки в узах и ребрах сети, долю потока дошедшего до моря, или до интересующего нас места (состояния, объявляемого поглощающим [18, 19]), доли конкретных типов расходов, потерь. На основании оценок по-

токов в узах сети легко оценить объемы воды поступающей в водохранилище или выходящей из водохранилища в течение заданного периода времени, уровень воды в водохранилищах.

Доля потока дошедшего до моря оценивается по формуле

$$P = \begin{pmatrix} E & 0 \\ R & Q \end{pmatrix} \quad \text{или} \quad P = \begin{pmatrix} Q & R \\ 0 & E \end{pmatrix},$$

где P - матрица переходных вероятностей системы, Q - матрица, получаемая из матрицы переходных вероятностей системы P вычеркиванием строк и столбцов соответствующих частям системы, отождествленными с поглощающими состояниями; E - единичная матрица; R - вектора-столбцы вероятностей перехода из водоемов сети в водоемы отождествляемые с соответствующими поглощающими состояниями.

Матрица $N = (I-Q)^{-1}$ - фундаментальная матрица системы; $B = (b_{ij}) = NR$, где b_{ij} - вероятность попадания потока выходящего из водоема i в водоем j (доля потока попавшего из водоема i в водоем j); $H = (h_{ij}) = (N-I)N_{dq}^{-1}$, где h_{ij} - вероятность попадания потока (доля потока), выходящего из водоема i системы в водоем j; A_{dq} - матрица, полученная из матрицы A, заменой всех недиагональных элементов нулями.

Пример

Сеть состоит из пяти водоемов: i - водоем, узел сети. i = 1, 2, 3, 4, 5.

Узел 5 (последний водоем (море) объявляется поглощающим состоянием. Добавляется еще два фиктивных поглощающих состояния 6 и 7. Поглощающее состояние 6 характеризует потоки, используемые нужды водоснабжения. Поглощающее состояние 6 характеризует потерянные (испарившиеся, ушедшие в землю) потоки.

12, 13, 23, 25, 34, 35, 45 – водотоки, смежные дуги.

16, 26, 36, 46 – водотоки, смежные дуги, характеризующие потоки, используемые на нужды водоснабжения. 17, 27, 37, 47 – водотоки, смежные дуги, характеризующие потоки потерянные (испарившиеся, ушедшие в землю) потоки.

На узлы сети 1, 2, 3, 4 поступают внешние потоки воды v_1 , v_2, v_3 , v_4 и потоки из других узлов сети.

Пусть λ_i – суммарный поток, входящий в i-ый узел. Поток λ_i распределяются по всем водотокам, смежным дугам ij:

$$\lambda_1 = \lambda_{12} + \lambda_{33} + \lambda_{16} + \lambda_{17}\,; \quad \lambda_2 = \lambda_{23} + \lambda_{24} + \lambda_{25} + \lambda_{26} + \lambda_{27}\,;$$

$$\lambda_3 = \lambda_{34} + \lambda_{35} + \lambda_{36} + \lambda_{37}\,; \quad \lambda_4 = \lambda_{45} + \lambda_{46} + \lambda_{47}\,.$$

Тогда λ_{ij} – поток, входящий в водоток, смежную дугу ij. $\lambda_{ij} = p_{ij}\lambda_i$. $\sum_j p_{ij} = 1$.

Расходы воды на водотоке ij связан с забором её на нужды водоснабжения, на потери, а также на пополнение запасов ресурса *Vj*(t) в хранилище (водоем, озеро, ильмень), привязанном к данному водотоку.

Матрица переходных вероятностей, описывающая рассматриваемую сеть, имеет вид

$$P = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{pmatrix} 0 & p_{12} & p_{13} & 0 & 0 & p_{16} & p_{17} \\ 0 & 0 & p_{23} & p_{24} & p_{25} & p_{26} & p_{27} \\ 0 & 0 & 0 & p_{34} & p_{35} & p_{36} & p_{37} \\ 0 & 0 & 0 & 0 & p_{45} & p_{16} & p_{16} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Системы потоковых уравнений сетевого уровня $\lambda_i = \sum_{j=1}^{7} p_{ij}\lambda_j + v_i$.

Решив данную систему, получим значения потоков входящих в узлы, характеризующие водоемы и потоки, используемые нужды водоснабжения и потерянные (испа-

ривщиеся, ушедшие в землю) потоки.

Доля потока дошедшего до моря, или до интересующего нас водоема, доли конкретных типов расходов, потерь, оцениваются формулами конечных цепей Маркова с поглощающими состояниями $P = \begin{pmatrix} Q & R \\ 0 & E \end{pmatrix}$, где P - матрица переходных вероятностей системы.

$$Q = \begin{pmatrix} 0 & p_{12} & p_{13} & 0 \\ 0 & 0 & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{34} \\ 0 & 0 & 0 & 0 \end{pmatrix},\ R = \begin{pmatrix} 0 & p_{16} & p_{17} \\ p_{25} & p_{26} & p_{27} \\ p_{35} & p_{36} & p_{37} \\ p_{45} & p_{46} & p_{47} \end{pmatrix},$$

$$N = (I - Q)^{-1} = \begin{pmatrix} 1 & -p_{12} & -p_{13} & 0 \\ 0 & 1 & -p_{23} & -p_{24} \\ 0 & 0 & 1 & -p_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1}.$$

Q - матрица, получаемая из матрицы переходных вероятностей системы P вычеркиванием строк и столбцов соответствующих частям системы, отождествленными с поглощающими состояниями;

Доля потока дошедшего до моря, или до интересующего нас водоема оценивается по формуле B = (b_{ij}) = NR, где b_{ij} - вероятность попадания потока выходящего из водоема i в водоем j (доля потока попавшего из водоема i в водоем j).

На основании оценок потоков в узах сети путем интегрирования оцениваются объемы воды поступающей в водохранилище или выходящей из водохранилища в течение заданного периода времени, объемы воды в водохранилищах в текущий момент времени.

На основании оценок объемы воды в текущий момент времени оцениваются уровни воды в водохранилищах.

Тем самым аппарат конечных цепей Маркова позво-

ляет упростить вычисление значений критериев оптимизации (объем вычислений не превосходит n^3) и значительно ускорить сам процесс оптимизации.

Заключение

Описываемые в статье сетевые модели дают возможность оценить:

1. Потоки в узах и ребрах сети как решение системы потоковых уравнений.
2. Долю потока дошедшего до моря, или до интересующего нас места, доли конкретных типов расходов, потерь используя формулы конечных цепей Маркова с поглощающими состояниями.
3. На основании оценок потоков в узах сети оцениваются объемы воды поступающей в водохранилище или выходящей из водохранилища в течение заданного периода времени, объемы воды в водохранилищах в текущий момент времени, уровни воды в водохранилищах.
4. Тем самым аппарат конечных цепей Маркова позволяет упростить вычисление значений критериев оптимизации и значительно ускорить сам процесс оптимизации, поскольку объем вычислений не превосходит n^3.

Список литературы

1. Водный объект. http://ru.wikipedia.org/wiki/%D0%92%D0%BE%D0%B4%D0%BD%D1%8B%D0%B9_%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82
2. Гидрографи́ческая сеть. http://ru.wikipedia.org/wiki/%D0%93%D0%B8%D0%B4%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D0%B5%D1%82%D1%8C
3. Прито́к. http://ru.wikipedia.org/wiki/%D0%9F%D1%80%D0%B8%D1%82%D0%BE%D0%BA

4. Рука́в реки.
http://ru.wikipedia.org/wiki/%D0%A0%D1%83%D0%BA%D0%B0%D0%B2_%D1%80%D0%B5%D0%BA%D0%B8
5. Бассе́йн водоёма.
http://ru.wikipedia.org/wiki/%D0%91%D0%B0%D1%81%D1%81%D0%B5%D0%B9%D0%BD_%D0%B2%D0%BE%D0%B4%D0%BE%D1%91%D0%BC%D0%B0
6. Уровень воды.
http://ru.wikipedia.org/wiki/%D0%A3%D1%80%D0%BE%D0%B2%D0%B5%D0%BD%D1%8C_%D0%B2%D0%BE%D0%B4%D1%8B
7. Расход воды.
http://ru.wikipedia.org/wiki/%D0%A0%D0%B0%D1%81%D1%85%D0%BE%D0%B4_%D0%B2%D0%BE%D0%B4%D1%8B
8. Гидрологи́ческий режи́м.
http://ru.wikipedia.org/wiki/%D0%93%D0%B8%D0%B4%D1%80%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B9_%D1%80%D0%B5%D0%B6%D0%B8%D0%BC
9. Гидрология.
http://ru.wikipedia.org/wiki/%D0%93%D0%B8%D0%B4%D1%80%D0%BE%D0%BB%D0%BE%D0%B3%D0%B8%D1%8F
10. Гидравлика.
http://ru.wikipedia.org/wiki/%D0%93%D0%B8%D0%B4%D1%80%D0%B0%D0%B2%D0%BB%D0%B8%D0%BA%D0%B0
11. Расчетная обеспеченность. Практическое пособие к СП 11-101-95.
http://lib.rushkolnik.ru/text/24651/
12. Триус Е.Б. Задачи математического программирования транспортного типа. - М.: Сов.радио, 1967. – 208 с.
13. Хранович И.Л. Потоковые модели. - М.: Научный мир, 2001. - 296 с.
14. Левит-Гуревич Л.К. Метод гидравлирасчета в естест-

венном русле с целью выбора противопаводковых мероприятий // Водные пути и русловые процессы. Труды Академии водохозяйственных наук. 1996. Вып. 3. - С. 71-77.
15. Левит-Гуревич Л.К. Метод динамического программирования для выбора рационального водораспределения в дельте реки // Известия Самарского НЦ РАН. 2010. Т. 12, № 1(4).
16. Гуревич И. М. Многоуровневая модель сети связи // Сб. Вопросы кибернетики. Протоколы и методы коммутации в вычислительных сетях. – М.: АН СССР, 1986.
17. Гуревич И. М. Динамическая модель сети связи. – В сб. Теория телетрафика в системах информатики. – М.: Наука, 1989.
18. Гуревич И.М. Исследование характеристик и свойств сетей массового обслу-живания. Самосогласованная система моделей // Материалы междунар. научной конф. «Современные вероятностные методы анализа, проектирования и автоматизации инфор-мационно-телекоммуникационных сетей». – Минск: Изд. центр БГУ. 2013. – С. 39-53.
19. Кемени Дж., Снелл Дж. Конечные цепи Маркова. М., «Мир», 1970.
20. Королюк В.С., Турбин А.Ф. Полумарковские процессы и их приложения. «Наукова думка», Киев -1976. - 184 с.

Научное издание

3-й международный
научно-технический
семинар
Материалы семинара.

СОВРЕМЕННЫЕ ПРОБЛЕМЫ
ПРИКЛАДНОЙ МАТЕМАТИКИ,
ИНФОРМАТИКИ, АВТОМАТИЗАЦИИ,
УПРАВЛЕНИЯ

9-13 сентября 2013 г., г. Севастополь

Технический редактор д.т.н., профессор Н.Б.Филимонов

Оригинал макет подготовлен в ИПИ РАН Юдичевой А.Л

Подписано в печать 17.07.2013
Тираж 70 экз.
Заказ 13-04

Издательство ИПИ РАН
119333, Москва, ул. Вавилова, д. 44, корп. 2

313