ISSN 1684–8853

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

6(49)/2010

НАУЧНЫЙ ЖУРНАЛ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

6(49)/2010

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

Адрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-44
Факс: (812) 494-70-18
E-mail: 80x@mail.ru
Сайт: www.i-us.ru

Журнал зарегистрирован в Министерстве РФ по делам печати, телерадиовещания и средств массовых коммуникаций. Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.

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

Журнал распространяется по подписке. Подписку можно оформить через редакцию, а также в любом отделении связи по каталогам: «Роспечать»: № 48060, № 15385; «Пресса России»: № 42476.

ЛР № 010292 от 18.08.98.
Сдано в набор 15.11.10. Подписано в печать 24.12.10. Формат 60×84 1/8.
Бумага офсетная. Гарнитура SchoolBookC. Печать офсетная.
Усл. печ. л. 11,6. Уч.-изд. л. 14,9. Тираж 1000 экз. Заказ 517.
Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.
Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.

УДК 004.896

СИСТЕМА ИНТЕЛЛЕКТУАЛЬНОГО УПРАВЛЕНИЯ МОБИЛЬНЫМ ИНФОРМАЦИОННО-СПРАВОЧНЫМ РОБОТОМ

М. В. Прищепа,
аспирантка
В. Ю. Будков,
аспирант
А. Л. Ронжин,
доктор техн. наук, доцент
Санкт-Петербургский институт информатики и автоматизации РАН

Проанализирован круг проблем, возникающих при разработке обслуживающих информационных роботов. Предложена модель интеллектуального управления мобильной информационной системой на базе многомодального интерфейса, обеспечивающего естественное человекомашинное взаимодействие.

Ключевые слова *— робототехника, системы интеллектуального управления, человекомашинное взаимодействие, мобильные подвижные системы.*

Введение

В настоящее время наибольшее распространение получили промышленные роботы, которые применяются на заводах, фабриках и иных производствах. Также разрабатываются различные военные роботы, большинство из которых представляют собой беспилотные летательные, подводные и наземные аппараты. Кроме того, особое внимание ученых и инженеров направлено на развитие обслуживающих роботов, предоставляющих ассистивные, информационные, обучающие и развлекательные услуги, например роботы-няньки, роботы-уборщики и др. [1]. В данной статье описаны результаты исследования по разработке информационно-справочного мобильного робота с интеллектуальным управлением на основе многомодального пользовательского интерфейса, который обеспечивает естественное взаимодействие клиентов с информационной системой.

Данный класс роботов предназначен для использования в музеях, выставочных и торгово-развлекательных комплексах в качестве гида и для предоставления посетителям различной полезной информации, например о тематике выставки, ее участниках, плане помещений, а также оказывает помощь в поиске нужного места или объекта. Современные обслуживающие роботы должны обеспечивать пользователю возможность интуитивного управления, их функции должны быть простыми и очевидными настолько, чтобы человек мог управлять ими без знания специальных команд или принципа работы, кроме того, система должна быть робастна к ошибкам в действиях пользователя. Изучение различных комбинаций многомодальных интерфейсов помогает разрешить фундаментальные вопросы человекомашинного взаимодействия и способствует созданию новых прикладных моделей в области безопасности, медицины, робототехники, логистики и других научных направлений.

Анализ требований к мобильному информационному роботу

Несмотря на то что концепция «робота-помощника» сформировалась уже несколько десятилетий назад, широкого распространения на рынке обслуживающие роботы пока не получили [1]. Основными сдерживающими причинами являются высокая стоимость, сложность интеграции, минимизации и совмещения всех компонентов в одном мобильном комплексе. Существуют проблемы в организации самостоятельной работы таких систем и в увеличении срока их работоспособности в автономном режиме.

Также имеется ряд недостатков в способах взаимодействия между человеком и системой, системой и окружающей средой. В последнее деся-

тилетие за рубежом активно проводятся исследования и разработки принципиально нового поколения обслуживающих информационных роботов с многомодальным пользовательским интерфейсом [2]. Это системы массового обслуживания, которые могут автоматически определять присутствие пользователя-клиента и общаться с ним естественным образом. Ввод информации может осуществляться как путем нажатия кнопок или сенсорного экрана, так и голосом или даже жестами. Такая система, как правило, обладает знанием о своем пространственном положении, а также планировке здания и использует эти данные при указании пользователю интересующего направления.

При разработке мобильного «робота-помощника» должен быть учтен ряд важных аспектов [1]. Во-первых, следует учитывать реальные потребности пользователей. Во-вторых, разработки должны вестись сразу на трех уровнях: технологии, необходимые для базовой конфигурации системы и использующиеся во всех приложениях; контекстно-зависимые технологии, которые могут быть адаптированы к большинству потребностей пользователей; персонифицированные технологии, разработанные в соответствии с требованиями конкретных пользователей. В-третьих, предложенные решения должны отличаться гибкостью, простотой пользования, надежностью и способностью к самовосстановлению, равно как функциональностью и адаптивностью.

При сопоставлении различных функциональных возможностей предлагаемых сегодня решений следует различать развлекательные системы, системы-помощники для людей с ограниченными возможностями и информационные системы, предназначенные для оповещения населения и предоставления различного рода услуг [3, 4]. Развлекательные системы, такие как японский гуманоидный минироботi-SOBOT, имеют минимальный набор функций и простейший интерфейс, основанный на управлении с компьютера или дистанционного пульта. Голосовое управление такими системами заключается в стандартном наборе команд, заранее включенных производителем. Системы-помощники имеют более широкий набор функций и выполняются в различных вариантах в зависимости от прикладной области. Широкое распространение получили роботы-животные; одним из наиболее популярных является японский робот Paro, предназначенный для лечения умственно и физически неполноценных пациентов. Он оснащен различными датчиками, реагирующими на прикосновения, освещенность, температуру и положение человека. Paro различает голосовые команды, узнает хозяина и выражает свои эмоции посредством движений и различных звуков. Также существуют роботы гуманоидного типа (такие как REEM-B, R100, Domo, STAIR, Twendy-One и др.), предназначенные для выполнения роли сиделки для пожилых людей, способные переносить предметы, передвигаться по помещению, следовать за хозяином, а в случае надобности помогать хозяину подняться с кресла. Такие системы обладают достаточно развитым голосовым интерфейсом, что позволяет пользователям общаться с роботами и управлять ими вербально.

Информационно-справочные системы (такие как SuperDroid RP2W) в основном предназначены для предоставления сведений об услугах и товарах различного вида. Подобные системы обладают минимальным набором функций, главный упор в их конструктивной части сделан на визуализацию предоставляемой информации. Они оснащены мультимедийным оборудованием и способны выполнять базовый набор команд, предусмотренных производителем. Иногда они имеют голосовой интерфейс с ограниченным словарем, могут автономно передвигаться по помещению, составляя карту местности, а также обладают стандартным набором датчиков для передвижения или обнаружения пользователей. Существующие информационные роботы различаются не только функциональными возможностями, но и уровнем коммуникации, связи с внешним миром, осведомленности о предпочтениях посетителей и способности к самообучению. От простейших систем, оснащенных механизмами включения/выключения, делается переход к сложным комплексным системам, способным интерпретировать поведение людей и отвечать их запросам.

Программно-аппаратные средства информационного робота

Наличие многомодального пользовательского интерфейса является отличительной характеристикой разрабатываемого мобильного информационно-справочного робота. Разработанные ранее технологии обработки аудио- и видеосигналов были адаптированы для применения на подвижной платформе в условиях, приближенных к реальной эксплуатации. Следует отметить, что системы распознавания речи, а также слежения за объектами по аудиовизуальной активности должны быть робастны по отношению к фоновым помехам и нестационарным шумам. Наиболее важные технологии, которые применены в мобильном роботе, это: локализация источников звука, автоматическое распознавание речи, идентификация диктора по речи, определение положения и слежение за подвижным объектом и лицом человека, аудиовизуальный синтез русской

речи «виртуальная говорящая голова» [5, 6]. Интеграция указанных технологий в рамках единого программно-аппаратного комплекса, оснащенного подвижной платформой с сенсорными датчиками препятствий, позволяет роботу перемещаться в пространстве и вести вербальный диалог с посетителями.

Аппаратная часть информационного робота состоит из двух основных компонентов: подвижной платформы и информационной сенсорной стойки. Общий вид мобильного информационного робота представлен на рис. 1. Компоновочная схема шасси платформы состоит из двух ведущих и двух флюгерных колес, что позволяет роботу двигаться в двух направлениях (вперед-назад), а также разворачиваться вокруг своей оси. Ультразвуковые и инфракрасные датчики, расположенные по внешней окружности платформы, служат для предотвращения столкновений с препятствиями при передвижении робота. Система управления приводами колес и обработки показаний датчиков установлена на бортовом компьютере информационной стойки, которая также оснащена встроенным Wi-Fi адаптером, двумя сенсорными мониторами, двумя массивами микрофонов, четырьмя веб-камерами, динамиками. Бортовой компьютер и остальные устройства подключены через источник бесперебойного питания к собственному аккумулятору.

Рис. 1. *Общий вид информационного мобильного робота*

С помощью видеокамер робот может следить за движущимися объектами и находить лица людей. Массивы микрофонов используются для локализации источников звука, распознавания речи, а также для идентификации посетителей. Обрабатывая поступающие потоки данных от гетерогенных сенсоров, система управления вырабатывает команды исполнительным устройствам и приложениям, обеспечивающим информационную поддержку пользователей.

Особенности системы управления роботом в торговом центре

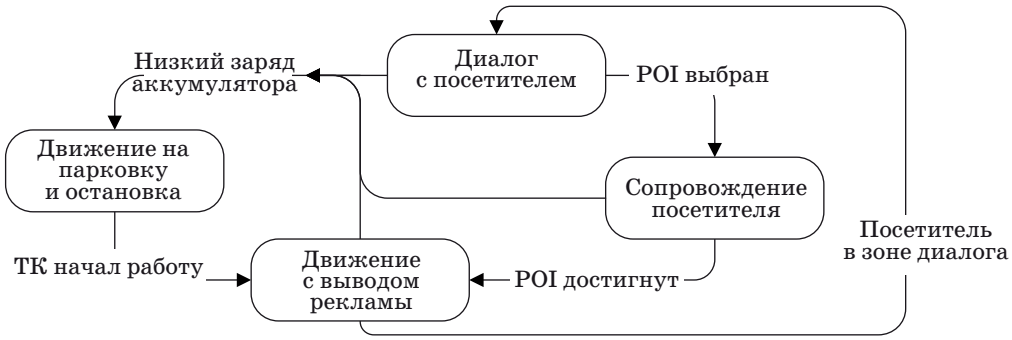
Одной из наиболее актуальных областей применения информационных роботов считаются торговые комплексы (ТК) [7]. Покупателям с каждым годом становится все сложнее ориентироваться в новых все расширяющихся комплексах, а непривычная новизна робота привлекает посетителей, поэтому его применение в рекламных целях весьма эффективно. Способность вести естественный диалог, перемещаться вместе с посетителем и учитывать его предпочтения является основным достоинством мобильного информационного робота.

Для формализации задачи интеллектуального управления информационным роботом была предложена концептуальная модель, включающая следующие основные сущности: зону обслуживания, объект, робота, посетителя, запрос, архив, режим работы, алгоритмическую базу (рис. 2). Предложенная модель служит основой для навигации робота в ТК, а также обработки голосовых запросов посетителей, интересующихся определенным товаром или магазином.

Основная информация, необходимая роботу о зоне обслуживания ТК, содержит карту допустимых маршрутов, место парковки, часы работы и список объектов: магазины, кинотеатры, кафе и другие точки интереса (*points-of-interest — POI*), в описании которых задаются их названия, расположение в комплексе, список предлагаемых услуг или товаров, рекламные ролики, параметры соединения с представителем объекта.

Множество режимов работы информационного робота включает: диалог с посетителем; сопровождение посетителя; движение с выводом рекламы; движение на парковку (рис. 3). В каждом из режимов рассчитывается свой маршрут передвижения и способ взаимодействия с посетителями. Также на выбор режима и изменение маршрута влияет возникновение динамических препятствий и состояние аккумуляторных батарей робота. При появлении посетителя в зоне речевого диалога производится аудиовизуальный синтез приветствия и запрашивается название инте-

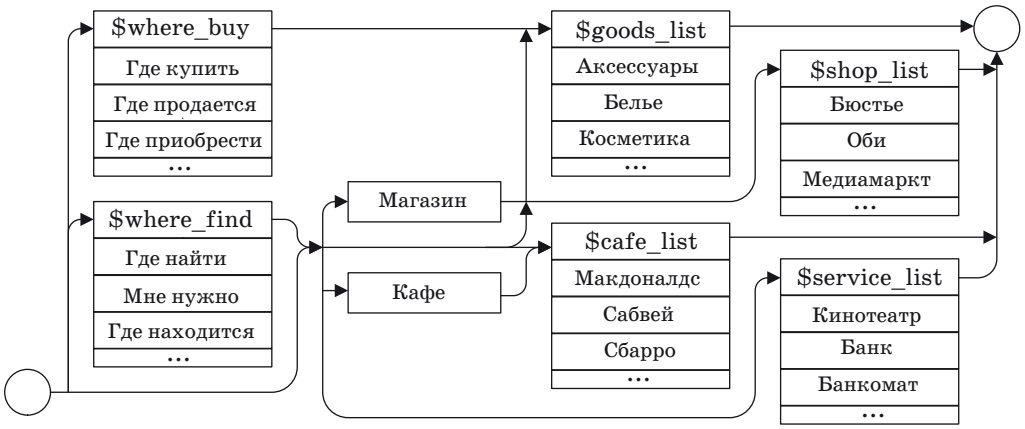
■ ***Рис. 2.*** *Концептуальная модель системы управления информационным роботом*



■ ***Рис. 3.*** *Логическая модель выбора режимов работы информационного робота*

ресующего POI, после чего робот сопровождает посетителя до нужного места и вновь переходит в режим рекламирования. При составлении маршрутов и модели диалога была проанализирована выложенная в Интернет карта торгового комплекса «МЕГА» в Санкт-Петербурге.

В ходе взаимодействия с посетителем формируется его профиль, включающий основные внешние данные и предпочтения в выборе товаров и магазинов. В разработанной модели диалога с посетителем основной целью является определение наименования POI или названия товара, поэтому структура допустимых фраз была составлена в виде грамматики, представленной на рис. 4. Фраза может содержать только название POI (элементы множеств $shop_list, $cafe_list,



■ ***Рис. 4.*** *Схема распознавания ключевых фраз пользователей информационного робота*

$service_list), название товара (элементы множества $goods_list) или название с дополнительными словами «магазин», «кафе», а также вступительными оборотами с глаголами «купить», «приобрести», «найти» и другими (элементы множеств $where_buy, $where_find).

Если пользователь во фразе указал название POI, то производится определение его местоположения, расчет маршрута и последующее движение робота к месту входа в интересующую POI. Если пользователь указал некоторый товар/услугу, то происходит поиск объектов, в которых продается данный товар, после чего список всех удовлетворяющих поиску объектов выводится на экран, где посетителю предлагается выбрать один. После выбора пользователем робот переходит в режим его сопровождения. Собранная в ходе движения и взаимодействия робота с посетителями информация о проделанных маршрутах, выполненных заданиях, предпочтениях пользователей архивируется для дальнейшей оптимизации системы управления роботом.

Заключение

Проектирование мобильных систем-помощников является новой научной парадигмой в области информационных технологий и актуальной научно-практической задачей. Разрабатываемый информационно-справочный робот представляет собой мобильную подвижную систему, которая содержит сеть интеллектуальных аппаратно-программных модулей, активационных устройств, мультимедийных средств и аудиовизуальных сенсоров. Основная задача робота — обеспечение пользователей необходимой справочной информацией и сервисами на основе автоматического анализа окружающей обстановки. Осведомленность системы о предпочтениях пользователя помогает более точно предсказать намерения и потребности человека. Наличие пользовательского многомодального интерфейса позволяет перевести взаимодействие человека с машиной на новый уровень и обеспечить понятный и доступный интерфейс любой категории пользователей.

Анализ реальных диалогов, способов взаимодействия с роботом и учет персонифицированных данных позволят выявить шаблоны поведения и предпочтений основных групп пользователей, сценариев человекомашинных взаимодействий и наиболее актуальные команды, которые следует генерировать и выполнять автоматически, что будет способствовать повышению эффективности взаимодействия человека с мобильным роботом.

Работа выполнена в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» (ГК № П876), а также поддержана КНВШ Администрации Санкт-Петербурга.

Литература

1. **Breazeal C.** et al. Humanoid robots as cooperative partners for people // Intern. J. of Humanoid Robots. 2004. Vol. 1. N 2. P. 1–34.
2. **Fritsch J.** et al. Multi-modal anchoring for human-robot-interaction // Robotics and Autonomous Systems. 2003. Vol. 43. P. 133–147.
3. **Fong T., Nourbakhsh I., Dautenhahn K.** A survey of socially interactive robots // Robotics and Autonomous Systems. 2003. Vol. 42. P. 143–166.
4. **Green S., Billinghurst X., Chen M., Chase G.** Human-robot collaboration: A literature review and augmented reality approach in design // Intern. J. of Advanced Robotic Systems. 2008. Vol. 5. N 1. P. 1–18.
5. **Ронжин А. Л., Карпов А. А., Кагиров И. А.** Особенности дистанционной записи и обработки речи в автоматах самообслуживания // Информационно-управляющие системы. 2009. № 5. С. 32–38.
6. **Карпов А. А., Ронжин А. Л.** Многомодальные интерфейсы в автоматизированных системах управления // Изв. вузов. Приборостроение. 2005. Т. 48. № 7. С. 9–14.
7. **Kanda T.** et al. An Affective Guide Robot in a Shopping Mall // Human-Robot Interaction 2009: Proc. of 4th ACM/IEEE Intern. Conf., Mar. 11–13, 2009, San Diego, USA. P. 173–180.

УДК 519.8

МНОГОКРИТЕРИАЛЬНЫЙ АНАЛИЗ ВЛИЯНИЯ ОТДЕЛЬНЫХ ЭЛЕМЕНТОВ НА РАБОТОСПОСОБНОСТЬ СЛОЖНОЙ СИСТЕМЫ

В. А. Зеленцов,
доктор техн. наук, профессор
А. Н. Павлов,
канд. техн. наук, доцент
Санкт-Петербургский институт информатики и автоматизации РАН

Предлагается метод решения задачи многокритериального анализа критичности отказов элементов сложной системы, основанный на комбинированном использовании метода нечеткого логического вывода и методов теории планирования эксперимента. Результирующий показатель критичности отказа элемента представляется в виде полинома, учитывающего влияние как отдельно взятых показателей, так и их совокупностей (по два, три и т. д.).

Ключевые слова *— геном структуры, критичность отказа, сложный объект, многокритериальный анализ, теория планирования эксперимента, лингвистические переменные.*

Введение

Современные технические системы и объекты содержат, как правило, большое число элементов. В этих условиях обеспечить требуемые характеристики работоспособности системы путем улучшения качества одновременно всех элементов вряд ли возможно, прежде всего, по экономическим причинам. Однако очевидно, что различные элементы в системе играют далеко не одинаковые роли, их отказы могут приводить к разным по степени влияния на состояние системы последствиям. Поэтому естественным является стремление сосредоточить усилия на совершенствовании элементов, играющих в обеспечении работоспособности системы наиболее важную роль. С целью выявить роль конкретных элементов (и их различных комбинаций) в обеспечении работоспособности всей системы применяются специальные показатели. Наиболее широко распространены два из них – структурная важность элемента и критичность отказов элемента [1, 2]. Но они не могут быть полностью определены только свойствами элемента и должны определяться в рамках сложной системы (СС), содержащей данный элемент.

В статье рассматриваются новые подходы к оцениванию данных показателей.

Показатель структурной важности элемента

К основным методам повышения надежности современных сложных систем относятся:

— технические методы, предполагающие улучшение надежности элементов СС;

— структурные методы, предполагающие изменение структурно-логического построения СС.

В состав современных СС, как правило, входят высоконадежные элементы. Показатель надежности всей системы практически нечувствителен к изменению показателей надежности отдельных элементов в реализуемом диапазоне значений. Это обуславливает важность использования структурных методов анализа влияния отказов элементов на работоспособность системы. Для исследования надежности СС, связанной с ее структурным построением (структурной надежности), используют структурные функции: надежности, безопасности, живучести, работоспособности, минимальных сечений отказов — путем ортогонализации монотонных и немонотонных функций алгебры логики, замещения логических аргументов в этих функциях вероятностями их истинности и соответствующих логических операций арифметическими.

В работах [3–5] введено понятие генома структуры, представляющего собой вектор $\chi = (\chi_0, \chi_1, \chi_2, ..., \chi_n)$, компонентами которого являются коэффициенты полинома функции минимальных сечений отказов структуры, составленной из однородных элементов. Помимо того, что геном структуры содержит информацию о топологических свойствах СС, с его помощью можно вычислять интегральные оценки значимости и вкладов отдельных элементов в структурную надежность СС с использованием вероятностного и нечетко-возможностного подходов, причем как для монотонных, так и для немонотонных структур. Вычисляемые характеристики вкладов элементов СС в структурный отказ (надежность) имеют самостоятельное значение и, кроме того, их можно рассматривать как один из показателей критичности отказов элементов.

Показатели критичности отказов

Критичность отказов элемента СС является векторным свойством, для оценивания которого используется целый ряд частных показателей, таких как [2]: степень тяжести последствий отказа; вероятность отказа; устойчивость элемента к воздействию внешних неблагоприятных факторов; степень резервирования; контролируемость состояния элемента; продолжительность существования риска отказа; возможность локализации отказа, а также рассмотренный выше показатель структурной надежности.

Перечисленные показатели критичности отказов могут иметь как количественный, так и качественный характер, и для их измерения могут использоваться различные виды шкал [2].

В самом общем случае критичность отказа элементов СС оценивается набором показателей $F = \{f_i, i = 1, ...,m\}$, каждый из которых представляет собой лингвистическую переменную и термы которых могут задаваться интервалами, нечеткими числами и т. п. Проведенные исследования [3–5] структурной надежности элементов показывают, что вклады элементов СС в структурную надежность (отказ) также представляют собой интервальные оценки.

Пример лингвистической шкалы применительно к одному из частных показателей приведен в табл. 1.

■ ***Таблица 1****. Лингвистическая шкала показателя*

Показатель	Шкала	Терм
Контроль состояния элемента	1. Состояние элемента не контролируется	1. Низкий
	2. Осуществляется периодический контроль	2. Ниже среднего
	3. Осуществляется постоянный контроль без прогнозирования	3. Средний
	4. Осуществляется периодический контроль с прогнозированием	4. Выше среднего
	5. Осуществляется постоянный контроль с прогнозированием	5. Высокий

Способ разрешения многокритериальной неопределенности при анализе критичности отказов

Выявление критичных элементов на основе их ранжирования по степени критичности отказов представляет собой задачу многокритериального выбора. Для разрешения многокритериальной неопределенности разработаны различные методы [6–10], обычно связанные со скаляризацией векторного критерия выбора за счет использования сверток различного вида. Использованием сверток обусловлены и основные недостатки данных методов:

— определение используемых в свертках весовых коэффициентов отдельных показателей сопряжено с серьезными трудностями получения и обработки экспертной информации, в результате весовые коэффициенты слабо связаны с действительной ролью частных показателей критичности при обобщенной оценке свойства критичности элементов объекта;

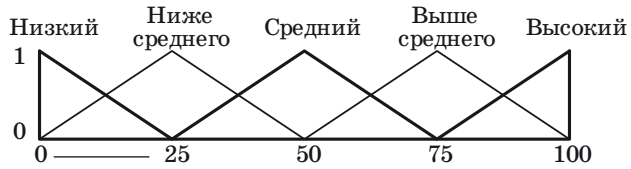
— не учитывается нелинейный характер влияния показателей друг на друга и на обобщенный показатель критичности отказа элементов СС;

— при построении интегрального показателя происходит выравнивание (сглаживание) значений частных показателей критичности элементов.

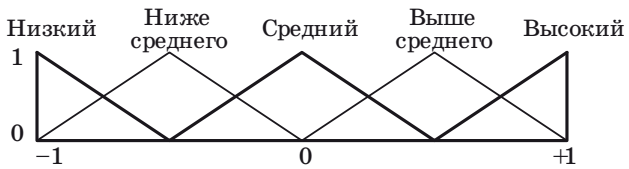
В данной статье предлагается комбинированный метод решения задачи многокритериального оценивания критичности отказов элементов СС, свободный от перечисленных недостатков. Он базируется на использовании метода нечеткого логического вывода [11, 12] и метода теории планирования эксперимента [12–15].

Сущность предлагаемого метода анализа критичности отказов

Рассмотрим лингвистическую переменную f_i = «Контроль состояния элемента СС». Она может принимать значения из множества простых и составных термов $T(f_i) = \{$«низкий», «ниже среднего», «средний», «выше среднего», «высокий»$\}$ (см. табл. 1).



■ ***Рис. 1.*** *Значения показателя «Контроль состояния элемента СС»*



■ ***Рис. 2.*** *Кодирование показателя «Контроль состояния элемента СС»*

Для формального представления термов лингвистических переменных можно использовать нечеткие числа (L-R)-типа. Тогда значения показателя «Контроль состояния элемента СС» можно представить по некоторой 100-балльной шкале (рис. 1).

Аналогично можно описать возможные значения других частных показателей.

Обобщенный взгляд лица, принимающего решение, на оцениваемую критичность отказа элемента СС формируется на основе анализа одновременно нескольких показателей с соответствующими значениями термов.

Введем для результирующего показателя лингвистическую переменную «Критичность отказа элемента СС», которая может принимать следующие значения: $T(f_{\text{рез}}) = \{$«низкая», «ниже среднего», «средняя», «выше среднего», «высокая»$\}$. Мнения экспертов о влиянии частных показателей критичности отказа элемента на результирующую оценку критичности в общем виде описываются следующими продукционными правилами:

P_j: «Если $f_1 = A_{1j}$ и $f_2 = A_{2j}$ и ... и $f_m = A_{mj}$, то $f_{\text{рез}} = A_{\text{рез}j}$», где $A_{ij} \in T(f_i)$, $A_{\text{рез}j} \in T(f_{\text{рез}})$.

Результирующий показатель критичности отказа элемента можно представить в виде полинома

$$f_{\text{рез}} = \lambda_0 + \sum_{i=1}^{m} \lambda_i f_i + \sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_{ij} f_i f_j + \ldots + \lambda_{12\ldots m} f_1 f_2 \ldots f_m,$$

учитывающего влияние как отдельно взятых показателей (через значения коэффициентов λ_i), так и совокупностей по два (λ_{ij}), три (λ_{ijk}) и т. д. показателей.

Для построения результирующего показателя необходимо перевести значения всех частных показателей f_i в шкалу $[-1, +1]$. С этой целью возможные крайние значения лингвистической переменной f_i маркируют как -1 и $+1$, при этом точка «0» соответствует середине шкалы (в соответствии с физическим смыслом данного показателя). Кодирование текущего значения лингвистической переменной f_i осуществляется по формулам

$$\bar{f}_i = (f_i - f_{\text{ср}}) / h,$$

где f_i — значение показателя на шкале лингвистической переменной; $f_{\text{ср}} = (f_{i\max} + f_{i\min}) / 2$ — средняя точка шкалы переменной; $h = (f_{i\max} - f_{i\min}) / 2$ — интервал варьирования; $f_{i\max}$, $f_{i\min}$ — крайние значения переменной. Результат кодирования представлен на рис. 2.

Далее необходимо построить матрицу опроса на профессиональном языке эксперта в крайних значениях показателей f_i. Матрица опроса для случая $m = 3$ представлена в табл. 2.

Так, например, во второй строке таблицы представлено следующее суждение эксперта: «Если показатель f_1 имеет значение “высокий”, показатель f_2 имеет значение “низкий”, показатель f_3 имеет значение “низкий”, то результирующий показатель $f_{\text{рез}}$ оценивается как “ниже среднего”».

Затем формируется ортогональный план экспертного опроса [11–15], который для случая m=3 представлен в табл. 3.

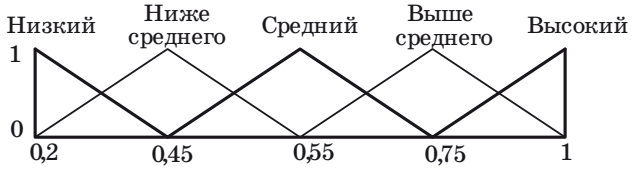
Пусть $f_{\text{рез}}$ может принимать значения, представленные на рис. 3.

■ ***Таблица 2.*** *Матрица опроса эксперта*

Высказывание	f_1	f_2	f_3	$f_{\text{рез}}$
1	Низкий	Низкий	Низкий	Низкий (Н)
2	Высокий	Низкий	Низкий	Ниже среднего (НС)
3	Низкий	Высокий	Низкий	Низкий (Н)
4	Высокий	Высокий	Низкий	Средний (С)
5	Низкий	Низкий	Высокий	Ниже среднего (НС)
6	Высокий	Низкий	Высокий	Выше среднего (ВС)
7	Низкий	Высокий	Высокий	Средний (С)
8	Высокий	Высокий	Высокий	Высокий (В)

■ ***Таблица 3.*** *Ортогональный план экспертного опроса*

	f_0	f_1	f_2	f_3	f_1f_2	f_1f_3	f_2f_3	$f_1f_2f_3$	$f_{рез}$
1	1	–1	–1	–1	1	1	1	–1	Н
2	1	1	–1	–1	–1	–1	1	1	НС
3	1	–1	1	–1	–1	1	–1	1	Н
4	1	1	1	–1	1	–1	–1	–1	С
5	1	–1	–1	1	1	–1	–1	1	НС
6	1	1	–1	1	–1	1	–1	–1	ВС
7	1	–1	1	1	–1	–1	1	–1	С
8	1	1	1	1	1	1	1	1	В



■ ***Рис. 3.*** *Значения интегрального показателя*

Для построения интегрального показателя с вещественными коэффициентами проведем операцию дефаззификации значений лингвистической переменной $f_{рез}$, для чего каждому терму поставим в соответствие моду его нечеткого числа («низкая» — 0,2; «ниже среднего» — 0,45; «средняя» — 0,55; «выше среднего» — 0,75; «высокая» — 1).

Расчет коэффициентов полинома производится по правилам, принятым в теории планирования эксперимента [11–15], для чего вычисляются усредненные скалярные произведения соответствующих столбцов ортогональной матрицы на вектор дефаззифицируемых значений результирующего показателя. Полученные результаты представлены в табл. 4.

Таким образом, свертка показателей в нашем случае имеет следующий вид:

$$f_{рез} = 0{,}5125+0{,}175f_1+0{,}0625f_2+0{,}1625f_3+ \\ +0{,}025f_1f_2+0{,}025f_1f_3+0{,}0375f_2f_3.$$

Если же не проводить дефаззификацию значений результирующего показателя, приведенная методика позволяет осуществлять построение функциональной зависимости интегрального показателя критичности отказов от f_i с нечеткими коэффициентами λ_0, λ_1, λ_2, ..., $\lambda_{12...m}$. Однако в данном случае следует воспользоваться арифметическими операциями над нечеткими трапецеидальными числами, введенными в работе [11].

Иллюстративный пример

Для исследования предложенного подхода рассмотрим небольшой пример [2] и сравним результаты ранжирования элементов СС по степени критичности предлагаемым методом и методом выделения паретовских слоев [2].

В качестве вектора критичности выберем двухкомпонентный вектор $\mathbf{f} = (f_1, f_2)$. Требуется проранжировать 10 элементов $X = \{x_i, i = 1, ..., 10\}$, имеющих следующие оценки критичности: $\mathbf{f}(x_1) = (0{,}1; 0{,}1)$, $\mathbf{f}(x_2) = (0{,}6; 0{,}1)$, $\mathbf{f}(x_3) = (0{,}8; 0{,}1)$, $\mathbf{f}(x_4) = (1{,}0; 0{,}1)$, $\mathbf{f}(x_5) = (0{,}1; 0{,}4)$, $\mathbf{f}(x_6) = (0{,}4; 0{,}4)$, $\mathbf{f}(x_7) = (0{,}6; 0{,}4)$, $\mathbf{f}(x_8) = (0{,}1; 0{,}6)$, $\mathbf{f}(x_9) = (0{,}8; 0{,}6)$, $\mathbf{f}(x_{10}) = (0{,}1; 0{,}8)$. Результаты выделения паретовских слоев с использованием классического отношения доминирования по Парето на множестве $X = \{x_i, i = 1, ..., 10\}$ представлены на рис. 4.

Здесь паретовские слои состоят из следующих элементов:

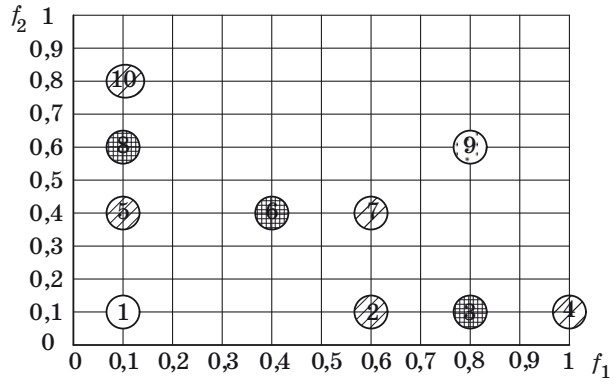
$$X_1^{nd} = \{x_1\},\ X_2^{nd} = \{x_2, x_5\},\ X_3^{nd} = \{x_3, x_6, x_8\},\\ X_4^{nd} = \{x_4, x_7, x_{10}\},\ X_5^{nd} = \{x_9\}.$$

Для ранжирования элементов по степени критичности в каждом слое использовалась [5] линейная свертка показателей вида $f_{рез}(x_i) = \lambda_1 f_1(x_i) + \lambda_2 f_2(x_i)$. Пусть коэффициенты важности показателей критичности одинаковы: $\lambda_1 = \lambda_2 = 0{,}5$. Тогда результат ранжирования элементов множества $X = \{x_i, i = 1, ..., 10\}$ следующий: $x_1 < x_5 < x_2 < x_8 < x_6 < x_3 < x_{10} < x_7 < x_4 < x_9$ (рис. 5).

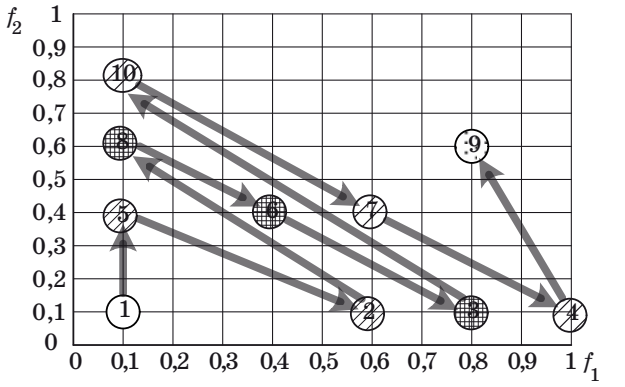
Несложно заметить, что в случае применения данного метода элементы x_6, x_7 в соответствующих паретовских слоях $X_3^{nd} = \{x_3, x_6, x_8\}$,

■ ***Таблица 4.*** *Результаты вычислений*

$f_0f_{рез}$	$f_1f_{рез}$	$f_2f_{рез}$	$f_3f_{рез}$	$f_1f_2f_{рез}$	$f_1f_3f_{рез}$	$f_2f_3f_{рез}$	$f_1f_2f_3f_{рез}$	Значения полинома
0,2	–0,2	–0,2	–0,2	0,2	0,2	0,2	–0,2	0,20
0,45	0,45	–0,45	–0,45	–0,45	–0,45	0,45	0,45	0,45
0,2	–0,2	0,2	–0,2	–0,2	0,2	–0,2	0,2	0,20
0,55	0,55	0,55	–0,55	0,55	–0,55	–0,55	–0,55	0,55
0,4	–0,4	–0,4	0,4	0,4	–0,4	–0,4	0,4	0,40
0,75	0,75	–0,75	0,75	–0,75	0,75	–0,75	–0,75	0,75
0,55	–0,55	0,55	0,55	–0,55	–0,55	0,55	–0,55	0,55
1	1	1	1	1	1	1	1	1,00
$\lambda_0 = 0{,}5125$	$\lambda_1 = 0{,}175$	$\lambda_2 = 0{,}0625$	$\lambda_3 = 0{,}1625$	$\lambda_{12} = 0{,}025$	$\lambda_{13} = 0{,}025$	$\lambda_{23} = 0{,}0375$	$\lambda_{123} = 0$	—



■ ***Рис. 4.*** *Разбиение элементов СС на паретовские слои*



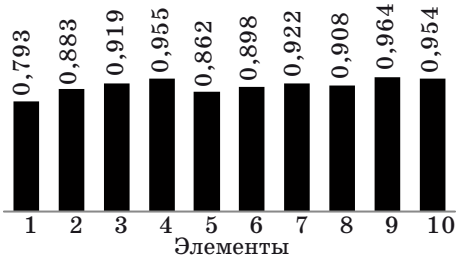
■ ***Рис. 5.*** *Результаты ранжирования элементов СС по методике работы* [*5*]

$X_4^{nd} = \{x_4, x_7, x_{10}\}$ при различных комбинациях коэффициентов важности λ_1, λ_2 в результате их ранжирования всегда будут доминироваться элементами данных слоев. Это означает, что при решении задач ранжирования и определения критичных элементов на паретовских слоях элементы x_6, x_7 ни при каких условиях не будут признаны наиболее или наименее критичными.

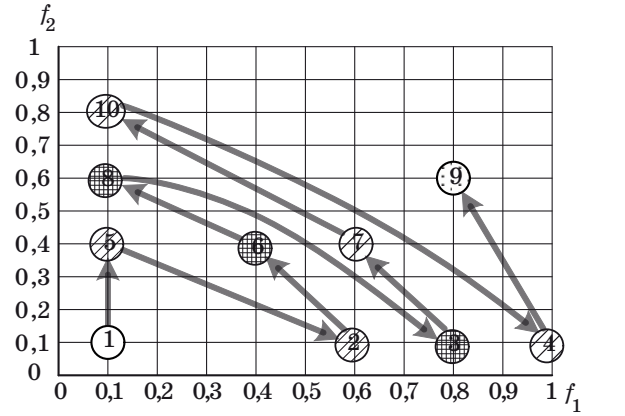
Проведем ранжирование элементов СС по степени их критичности предлагаемым комбинированным методом. Покажем, что из паретовских слоев можно выделять элементы x_6, x_7 так же, как и любые другие. Ортогональный план экспертного опроса представлен в табл. 5.

Обработав данные экспертного опроса по изложенной выше методике, получим следующее уравнение, описывающее мнение эксперта: $f_{рез} = 0{,}75+0{,}2f_1+0{,}25f_2-0{,}2f_1f_2$. Результаты вычисления интегрального показателя критичности отказов элементов показаны на рис. 6 и 7.

Полученные результаты ранжирования элементов СС по степени их критичности схожи с вышепредставленной ранжировкой (см. рис. 5). При этом элементы x_6, x_7 в соответствующих паретовских слоях $X_3^{nd} = \{x_3, x_6, x_8\}$, $X_4^{nd} = \{x_4,$



■ ***Рис. 6.*** *Значения интегрального показателя критичности отказов для элементов*



■ ***Рис. 7.*** *Результаты ранжирования элементов СС по предлагаемой методике*

■ ***Таблица 5.*** *Вычисление значений коэффициентов полинома*

f_0	f_1	f_2	$f_{рез}$	f_1f_2
1	–1	–1	0,1	1
1	1	–1	0,9	–1
1	–1	1	1	–1
1	1	1	1	1
$f_0f_{рез}$	$f_1f_{рез}$	$f_2f_{рез}$	$f_1f_2f_{рез}$	Значения полинома
0,1	–0,1	–0,1	0,1	0,100
0,9	0,9	–0,9	–0,9	0,900
1	–1	1	–1	1,000
1	1	1	1	1,000
$\lambda_0 = 0{,}75$	$\lambda_1 = 0{,}2$	$\lambda_2 = 0{,}25$	$\lambda_{12} = -0{,}2$	–

$x_7, x_{10}\}$, по мнению эксперта, оказались наименее критичными.

Заключение

Использование вместе с количественной также и качественной (нечеткой, неточной, интервальной) информации о влиянии отказов элементов на функционирование СС существенно повышает достоверность выводов и принимаемых решений при проектировании и управлении СС.

В рамках предложенного метода осуществляется формализация экспертной информации, представленной на естественном для эксперта языке, путем введения лингвистических переменных, которые позволяют адекватно отобразить приблизительное словесное описание предметов и явлений даже в тех случаях, когда детерминированное описание отсутствует или невозможно в принципе.

Предложенный метод анализа критичности отказов позволяет формализовать опыт эксперта (группы экспертов) в виде прогностических моделей в многомерном пространстве и учесть комплексное влияние одновременно нескольких факторов на результирующий показатель критичности СС. За счет этого выявляется нелинейный характер влияния частных показателей на интегральный показатель критичности отказов и повышается достоверность принимаемых решений.

Исследования проводились при финансовой поддержке РФФИ (гранты 08-08-00346, 08-08-00403, 09-08-00259, 10-08-90027, 10-07-05019), Отделения нанотехнологий и информационных технологий РАН (проект № О–2.3/03).

Литература

1. **Ефремов А. С., Зеленцов В. А., Миронов А. Н., Холоименко К. А.** Критерии предельного состояния координатных АТС // Вестник связи. 2004. № 2. С. 71–76.
2. **Афанасьев В. Г., Зеленцов В. А., Миронов А. Н.** Методы анализа надежности и критичности отказов сложных систем / МО. — СПб., 1992. — 99 с.
3. **Павлов А. Н., Соколов Б. В., Сорокин М. В.** Анализ структурной динамики комплексной системы защиты информации // Информация и безопасность. 2009. Т. 12. № 3. С. 389–396.
4. **Павлов А. Н.** Логико-вероятностный и нечетко-возможностный подходы к исследованию монотонных и немонотонных структур // Кибернетика и высокие технологии XXI века: Тез. докл. XI Междунар. науч.-техн. конф., Воронеж, 12–14 мая 2010 г. / НПФ «САКВОЕЕ». Воронеж, 2010. С. 483–492.
5. **Павлов А. Н.** Исследование немонотонных систем: анализ «мостиковой» структуры // Моделирование и анализ безопасности и риска в сложных системах: Тр. X Междунар. науч. школы МА БР-2010, Санкт-Петербург, 6–10 июля 2010 г. СПб.: ГУАП, 2010. С. 85–93.
6. **Соколов Б. В.** и др. Военная системотехника и системный анализ: учебник / МО. — СПб., 1999. — 408 с.
7. Нечеткие множества в моделях управления и искусственного интеллекта / Под ред. Д. А. Поспелова. — М.: Наука, 1986. — 312 с.
8. **Борисов А. Н., Крумберг О. А., Федоров И. П.** Принятие решений на основе нечетких моделей. — Рига: Зинанте, 1990. — 184 с.
9. **Павлов А. Н., Соколов Б. В.** Принятие решений в условиях нечеткой информации: учеб. пособие. — СПб.: ГУАП, 2006. — 72 с.
10. **Андрейчиков А. В., Андрейчикова О. Н.** Анализ, синтез, планирование решений в экономике: учебник. — М.: Финансы и статистика, 2000. — 368 с.
11. **Спесивцев А. В.** Управление рисками чрезвычайных ситуаций на основе формализации экспертной информации. — СПб.: Изд-во Политехн. ун-та, 2004. — 238 с.
12. **Павлов А. Н.** Методика построения псевдоуниверсальных сверток лингвистических показателей на основе теории планирования эксперимента: сб. докл. // XI Междунар. конф. по мягким вычислениям и измерениям (SCM'2008), РФ, Санкт-Петербург, 23–25 июня 2008 г. — СПб.: СПбГЭТУ «ЛЭТИ», 2008. Т. 1. С. 169–172.
13. **Налимов В. В., Чернова Н. А.** Статистические методы планирования экстремальных экспериментов. — М.: Наука, 1965. — 382 с.
14. **Налимов В. В.** Теория эксперимента. — М.: Наука, 1971. — 208 с.
15. **Адлер Ю. П., Маркова Е. В., Грановский Ю. В.** Планирование эксперимента при поиске оптимальных условий. — М.: Наука, 1976. — 280 с.

УДК 004.9

ГРАДИЕНТНЫЙ МЕТОД КООРДИНАЦИИ УПРАВЛЕНИЙ ИЕРАРХИЧЕСКИМИ И СЕТЕВЫМИ СТРУКТУРАМИ

А. Я. Фридман,
доктор техн. наук, профессор
О. В. Фридман,
канд. техн. наук
Институт информатики и математического моделирования технологических процессов Кольского научного центра РАН

Представлен градиентный метод координации децентрализованного управления иерархическими и сетевыми структурами на основе предложенных ранее необходимых и достаточных условий координируемости локально организованной иерархии динамических систем. Работоспособность метода проиллюстрирована результатами математического моделирования двухуровневой системы управления линейными объектами. Показано, что подключение локального управления и координации расширяет диапазон устойчивости системы к внешним и структурным возмущениям, а также повышает ее быстродействие в несколько раз.

Ключевые слова *— ситуационный анализ и синтез, концептуальная модель предметной области, координируемость управляемых систем.*

Введение

Возникновение иерархической структуры управления сложными объектами обусловлено возрастающей сложностью централизованного управления ими. Поэтому появилась необходимость разделения всего процесса принятия решений на такое число уровней, чтобы решение задачи оптимизации на каждом из них имело приемлемую сложность. Но с возникновением многоуровневых иерархических систем управления появилась и новая задача согласования и координации решений, принимаемых на всех уровнях управления [1].

Метод координации основан на предложенных [2, 3] необходимых и достаточных условиях координируемости локально организованной иерархии динамических систем. Компьютерный эксперимент проводился с помощью системы визуального блочного математического моделирования VisSim [4]. В качестве объекта координации вначале рассматривалась двухуровневая иерархическая система управления линейным объектом, затем полученные результаты были обобщены для объекта с сетевой структурой. Особенность подобных объектов по сравнению с иерархическими состоит в том, что в них все или большинство лиц, принимающих решения, имеют равный ранг. Пример — задачи формирования виртуальных предприятий [5]. Координация таких систем может производиться только на метауровне.

Цель анализа состояла в выявлении диапазонов устойчивости локальных управлений и координирующих сигналов к небольшим изменениям динамических характеристик объекта управления (вариациям матрицы динамики объекта). Кроме того, исследовались возможности повышения быстродействия децентрализованной системы управления.

Постановка задачи

В системе ситуационного моделирования (ССМ) [6] изучаемая динамическая система должна быть представлена в виде иерархически упорядоченного множества объектов (составных частей). Эта иерархия отражает организационные взаимоотношения объектов. Критерий качества работы каждого объекта имеет вид

$$\Phi ::= \left(\frac{1}{m} \sum_{i=1}^{m} \left(\frac{a_i - a_{i0}}{\Delta a_i} \right)^2 \right)^{1/2} ::= \left(\frac{1}{m} \sum_{i=1}^{m} \delta a_i{}^2 \right)^{1/2}, \quad (1)$$

где a_i — сигналы из списка выходных параметров данного объекта, их общее количество равно m; a_{i0} и $\Delta a_i > 0$ — настроечные параметры, отражающие требования вышестоящего объекта к номи-

нальному значению a_i и допустимому отклонению Δa_i от этого значения соответственно; $\delta a_i ::= \frac{a_i - a_{i0}}{\Delta a_i}$ — относительное отклонение фактического значения сигнала a_i от его номинального значения a_{i0}.

Если считать a_i скалярными критериями качества работы элемента модели, номинальные значения которых определяются величинами a_{i0}, то (1) представляет собой обобщенный критерий с коэффициентами важности, обратно пропорциональными допустимым отклонениям скалярных критериев, что не противоречит здравому смыслу.

Из (1) также следует, что в ССМ для координации применяется способ прогнозирования взаимодействий [1]. Глобальная задача ставится путем выбора доминирующего скалярного критерия, который должен вносить минимальный вклад в обобщенный критерий (1). Пусть для определенности это будет $a_{10}^{(0)}$.

Рассмотрим возможности применения критерия (1) для координации локальных управлений при управлении иерархическими и сетевыми объектами.

Исследование иерархической системы управления

Как и в работе [1], будем без потери общности рассматривать двухуровневую систему (рис. 1), в которой объект верхнего уровня (координатор) O_0, имеющий обобщенный критерий качества Φ_0 типа (1), передает подчиненным ему объектам (подобъектам) O_1–O_n, имеющим аналогичные критерии качества, настроечные параметры и получает в качестве сигналов обратной связи относительные отклонения фактических значений локальных критериев качества подобъектов от их номинальных значений. Подобъекты взаимодействуют через управляемую систему и не имеют информации о состоянии других подобъектов, т. е. вся система локально организована.

Предлагаемый принцип координации такой системы с точки зрения системного анализа соответствует внешнему (объективному) подходу к оценке эффективности функционирования подсистем в составе метасистемы. Этот принцип состоит в следующем: задачи подобъектов будут скоординированы относительно задачи координатора, если знак градиента обобщенного критерия координатора по его текущему доминирующему скалярному критерию совпадает со знаками градиентов этого обобщенного критерия по всем текущим значениям скалярных критериев подобъектов.

Из (1) имеем

$$\frac{\partial \Phi_k}{\partial a_i^{(k)}} = \frac{2}{m_k} \frac{a_i^{(k)} - a_{i0}^{(k)}}{\Delta^2 a_i^{(k)}}, \qquad (2)$$

откуда следует, что знак производной можно менять нужным образом, выбирая величину $a_{i0}^{(k)}$ больше или меньше $a_i^{(k)}$. С другой стороны, если считать, что действия всех подобъектов равно важны для достижения цели координатора (возможность обобщения очевидна), то

$$\frac{\partial \Phi_0}{\partial a_i^{(k)}} = \sum_{j=1}^{m_0} \frac{\partial \Phi_0}{\partial a_j^{(0)}} \frac{\partial a_j^{(0)}}{\partial a_i^{(0)}} = \frac{2}{m_0} \sum_{j=1}^{m_0} \mu_j \frac{\partial a_j^{(0)}}{\partial a_i^{(0)}} \approx$$

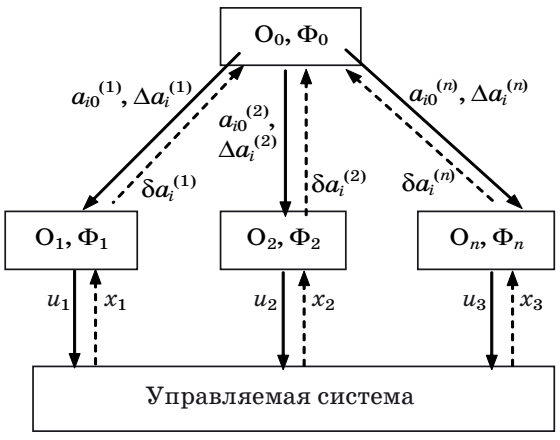
$$\approx \frac{2}{n m_0} \sum_{j=1}^{m_0} \mu_j \frac{\mathrm{Inc}[a_j^{(0)}]}{\mathrm{Inc}[a_i^{(0)}]}, \qquad (3)$$

где обозначено $\mu_j = \frac{a_j^{(0)} - a_{j0}^{(0)}}{\Delta^2 a_j^{(0)}}$, а Inc[*] есть приращение (инкремент) параметра в скобках за предыдущий временной шаг.

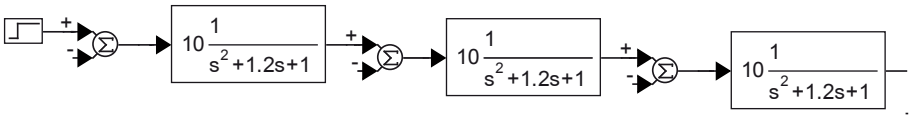
Система будет координируема, если координатор выберет все $a_{i0}^{(k)}$ таким образом, чтобы знаки величин (2) (для $k = 0$ и $i = 1$) и (3) (для всех k от 1 до n и всех i для каждого подобъекта) совпадали.

Полученные достаточные условия координируемости аналогичны идеям обеспечения устойчивости локального управления в коллективах автоматов [7], где требуется положительность частных производных обобщенного критерия типа (1) по входным параметрам соответствующего элемента коллектива.

В целях подтверждения теоретических результатов на математической модели исследовалась устойчивость характеристик децентрализованного управления на основе градиентов локальных критериев качества и возможности по-



■ ***Рис. 1.*** *Двухуровневая многоцелевая система*



■ ***Рис. 2.*** *Схема модели управляемой системы*

вышения (оптимизации) быстродействия децентрализованной системы.

Моделирование иерархической системы проводилось на примере управляемого объекта, представляющего собой три последовательно соединенных линейных звена с передаточной функцией второго порядка, одним управляющим входом и одним выходом каждый. Рассматривалась двухуровневая система управления (см. рис. 1).

В качестве управляемой системы при моделировании использовалась линейная трехблочная система (рис. 2).

Были построены три аналогичных друг другу управляющих элемента нижнего уровня, соответствующие элементам второго уровня на рис. 1. В каждом из них вычисляется градиент обобщенного критерия (2), его значения подаются в качестве управляющего воздействия на вход каждого из трех блоков управляемой системы.

Для принятия решений (выработки управляющих воздействий) управляющие элементы нижнего уровня использовали локальную информацию о состоянии подчиненных им звеньев управляемого объекта, координатор обладал полной информацией о состоянии этого объекта и управляющих элементов нижнего уровня, что соответствует принципам теории иерархических систем [1].

Проведенный модельный эксперимент включал несколько последовательных этапов. Первый этап состоял в исследовании устойчивости системы к малым возмущениям. На втором этапе эксперимента на блоки исследуемой системы подавалось управляющее воздействие, вычисляемое в соответствии с (2). Аналогично первому этапу исследований выявлялись диапазоны устойчивости системы при подключении управления на отдельный блок, попарно и на все три блока. Значения коэффициентов усиления при вводе управлений подбирались по значению установившейся погрешности реальной траектории относительно идеальной при условии сохранения устойчивости возмущенной системы. На следующем этапе моделирования подключался координатор (верхний уровень на рис. 1) и изменялись значения коэффициентов усиления приращений координирующих сигналов (номинальных значений a_{i0}) для повышения быстродействия децентрализованной системы. Последний этап моделирования состоял в выявлении диапазонов устойчивости системы при наличии управления и координации.

Результаты моделирования иерархической системы управления

Первый этап. В качестве возмущений рассматривались перекрестные связи между отдельными блоками управляемой системы, изменяющие собственные числа матрицы динамики системы. Возмущающий коэффициент K_{ij} обозначает подачу сигнала на вход i-го блока с выхода j-го блока. Таким образом, для трехблочной системы рассматриваются коэффициенты структурных возмущений K_{12}, K_{13}, K_{23}.

В ходе эксперимента были исследованы все возможные сочетания подключений возмущающих воздействий — по одному, попарно, все три одновременно. Эксперимент показал, что наиболее значимое воздействие на устойчивость системы оказывает изменение K_{13}, а наименьшее — изменение K_{12}. Кроме того, были выявлены диапазоны изменений коэффициентов, в пределах которых система оставалась устойчивой с заданной 5 %-й точностью.

Второй этап. На блоки исследуемой системы подавалось управляющее воздействие согласно (2). Аналогично первому этапу исследований выявлялись диапазоны устойчивости системы при подключении управления на отдельный блок, попарно и на все три блока. Получено, что наибольший эффект дает подключение всех трех блоков, причем подключение управления существенно расширяет диапазоны устойчивости (табл. 1).

Далее был осуществлен подбор значений коэффициентов усиления при вводе локальных управлений. Подбор проводился в условиях устойчивости возмущенной системы. Наилучшее быстродействие получено при коэффициенте

■ ***Таблица 1***

Одновременное подключение	K_{12}	K_{13}	K_{23}
Без управления	−0,0001 ÷ 0,0001	−0,00001 ÷ 0,00001	−0,0001 ÷ 0,0001
С управлением	−0,001 ÷ 0,001	−0,0001 ÷ 0,0001	−0,0005 ÷ 0,0005

0,608 для первого блока и коэффициенте 1 для второго и третьего блоков.

На ***третьем этапе*** подключались все блоки модели.

Блок координации, который соответствует верхнему блоку на рис. 1, содержит три одинаковых подблока. На вход каждого подблока координатора подается фактическое значение сигнала a_i, номинальное значение сигнала a_{i0} и рассчитывается относительное отклонение фактического значения сигнала a_i от его номинального значения: $a_{i0} - \delta a_i ::= \frac{a_i - a_{i0}}{\Delta a_i}$.

Далее вычисляется «новое» номинальное значение $a_{i0}' = a_{i0} + \Delta a_{i0}$, где $\Delta a_{i0} = k_i \delta a_i$. Значение коэффициента k_i изначально полагается равным единице. На следующем этапе оно изменялось для повышения быстродействия системы.

Эксперимент показал, что подключение координатора улучшает установившуюся погрешность в несколько раз, если оценивать ее по значению отклонения стабилизировавшихся сигналов друг от друга. Представленные графики соответствуют состояниям возмущенной системы без управления (рис. 3, *а*), с подключенным нижним уровнем управления (рис. 3, *б*) и подключенным управлением и координацией (рис. 3, *в*). Значения возмущающих коэффициентов во всех трех случаях не менялись: $K_{12} = 0{,}002$, $K_{13} = -0{,}0001$, $K_{23} = -0{,}002$. Наличие управления вдвое снижает процент расхождения траекторий эталонной и исследуемой систем. Подключение координатора позволяет повысить устойчивость системы к внешним возмущениям еще в два раза.

Далее решался вопрос повышения быстродействия всей системы.

Оказалось, что различные сочетания значений коэффициентов k_i для разных блоков координатора в существенно различной степени влияют на результат моделирования, в частности, на значение процента сходимости и время сходимости идеальной и реальной кривых. Установившаяся погрешность при подключенном управлении нижнего уровня без координатора — 8,41 %. Наилучшее быстродействие достигнуто при $k_1 = 5$, $k_2 = k_3 = -8000$; установившаяся погрешность составила 3,67 %.

Из графиков на рис. 3 видно, что при наличии управления и координации время сходимости идеальной и реальной кривых составляет примерно 10 с, тогда как без координации (или неоптимальных значениях коэффициентов k_i) при тех же возмущениях кривые вообще не сходились или при других значениях возмущений сходились примерно через 20 с. Таким образом, подключение блока координации повышает быстродействие системы приблизительно вдвое.

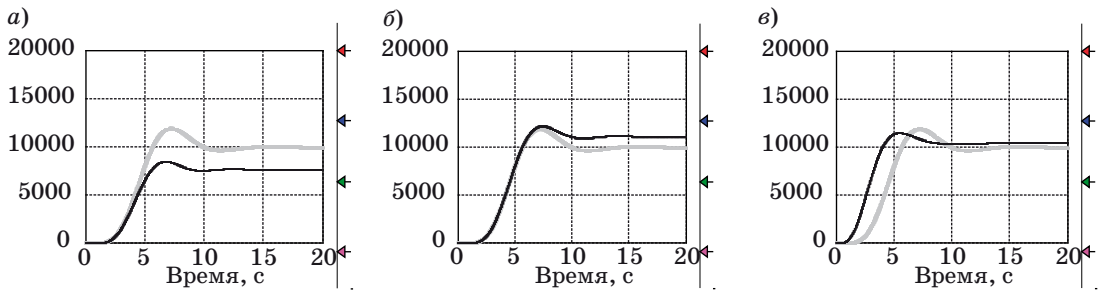
На последнем, ***четвертом этапе*** исследований иерархической системы выявлялись диапазоны устойчивости системы при наличии и управления, и координации аналогично тому, как это производилось на предыдущих этапах. Эксперимент показал, что диапазоны устойчивости системы существенно расширились по сравнению со случаем, когда подключалось только управление (см. табл. 1), и составили: $K_{12} = -0{,}002 \div 0{,}002$, $K_{23} = -0{,}002 \div 0{,}002$, $K_{13} = -0{,}00015 \div 0{,}00015$.

Исследование децентрализованной системы управления сетью объектов

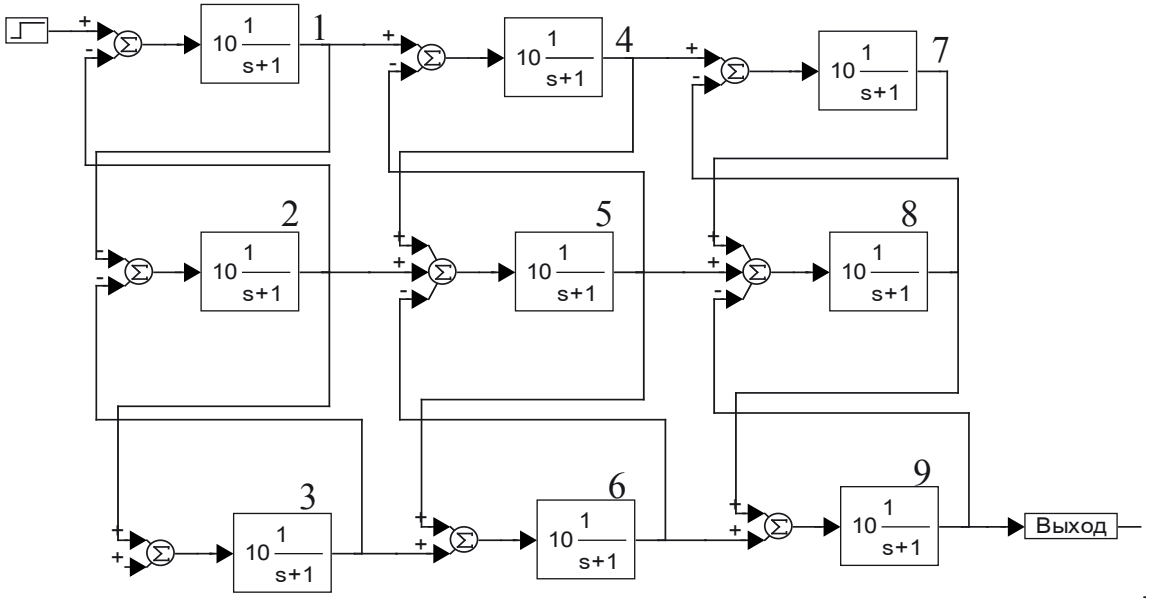
По методике, описанной для иерархической управляемой системы, были проведены также исследования децентрализованной системы управления сетью объектов. Ввиду усложнения модели принято решение об упрощении вида передаточных функций по сравнению с иерархической системой.

Схема эталонной сетевой структуры, на вход которой подается ступенчатый сигнал с амплитудой +10, показана на рис. 4.

Сначала была проанализирована устойчивость исследуемой системы к внешним возмуще-



■ ***Рис. 3.*** *Влияние управляющих и координирующих воздействий на устойчивость системы*: *а — без управления* (*установившаяся погрешность 23,1 %*); *б — с подключенным нижним уровнем управления* (*установившаяся погрешность 11,2 %*); *в — с управлением и координацией* (*установившаяся погрешность 4,63 %*): ——— *— идеальная*; ——— *— управляемая*



■ ***Рис. 4.*** *Схема эталонной сетевой структуры*

ниям. Для этого на каждый узел сети поочередно подавался сигнал, аналогичный входному, но с амплитудой +1, что соответствует 10 %-му внешнему возмущению.

Далее определялись диапазоны устойчивости системы к малым внутренним возмущениям, реализованным путем добавления обратных связей между выходами и входами узлов сети (в направлении от общего выхода системы к общему входу), по той же методике, что и для иерархической системы (отклонение по амплитуде ± 5%). Проанализированы все возможные сочетания связей «выход-вход».

Затем определялись диапазоны устойчивости системы при поочередном подключении управления на каждый узел сети. Управления задавались пропорционально градиенту обобщенного критерия (2), его значения подавались в качестве управления по одному на вход каждого из возбуждаемых узлов сети.

Исследовалось поведение системы при одновременном подключении всех управляющих элементов, возбуждение подавалось только на один узел сети. Определены диапазоны устойчивости для такой ситуации.

Следующим шагом эксперимента было подключение координатора, построенного аналогично иерархической системе. Исследовано подключение блока координации только на возбуждаемый узел и полное подключение координатора (на все узлы сети) с одиночным подключением управления (на возбуждаемый узел сети) и полным подключением управления (на все узлы сети).

Результаты моделирования сетевой системы управления

На внешнее возмущение реагировали только три первых блока, причем оно компенсировалось уже при одиночном подключении управляющего элемента на возбуждаемый узел сети. В целом сеть продемонстрировала устойчивость к воздействиям такого рода.

По величине диапазонов устойчивости обратные связи между узлами сети можно условно разбить на «сильные» и «слабые»; оказалось, что «сильные» связи замыкаются в основном на три первых узла сети. В табл. 2 показаны результаты исследований устойчивости сети к малым структурным возмущениям.

При поочередном подключении управления на каждый узел сети диапазоны устойчивости системы для «сильных» связей расширялись в среднем на порядок, а для «слабых» связей практически не менялись, но реакция «слабых» связей появлялась не только на возмущаемом узле сети, но и на узлах 4 и 8 независимо от того, на какой узел подавалось возмущение. Подключение одиночного управления на возбуждаемый узел сети достаточно эффективно компенсирует небольшие структурные возмущения.

При моделировании одновременного подключения всех управляющих элементов диапазоны устойчивости для «сильных» связей в среднем не изменились, для связей, замкнутых на первый узел сети, — незначительно расширились, а для других — уменьшились в 2–4 раза по сравнению с одиночным подключением управления. Диапа-

■ *Таблица 2*

«Сильные» связи		«Слабые» связи	
Выход-вход	Диапазон устойчивости	Выход-вход	Диапазон устойчивости
2–1	0,0003 ÷ −0,001	5–4	0,4 ÷ −0,25
3–1	0,000001 ÷ −0,000001	6–4	0,5 ÷ −0,5
3–2	0,00025 ÷ −0,00025	6–5	0,1 ÷ −0,1
4–1, 4–3, 6–3, 7–2, 9–2	0,000005 ÷ −0,000005	7–4	0,01 ÷ −0,05
4–2	0,00005 ÷ −0,00005	7–5	0,009 ÷ −0,009
5–1, 5–3	0,000025 ÷ −0,00002	8–2	0,002 ÷ −0,002
5–2	0,0001 ÷ −0,0001	8–4	0,005 ÷ −0,06
6–1	0,000001 ÷ −0,000005	8–5	0,07 ÷ −0,1
6–2	0,00005 ÷ −0,00004	8–6	0,064 ÷ −0,001
7–1, 7–3, 9–1, 9–3	0,0000005 ÷ −0,0000005	8–7	0,001 ÷ −0,001
7–6	0,00001 ÷ −0,00001	9–4	0,05 ÷ −0,05
8–1	0,0002 ÷ −0,0002	9–5	0,001 ÷ −0,02
8–3	0,0002 ÷ −0,0002	9–7	0,01 ÷ −0,009
9–6	0,0001 ÷ −0,0005	9–8	0,1 ÷ −0,15

зоны устойчивости «слабых» связей по-прежнему не менялись, но к узлам сети, всегда проявляющим реакцию на возмущение, кроме четвертого и восьмого узлов, добавились пятый и седьмой. Таким образом, локальные управления при отсутствии координации «мешали» друг другу.

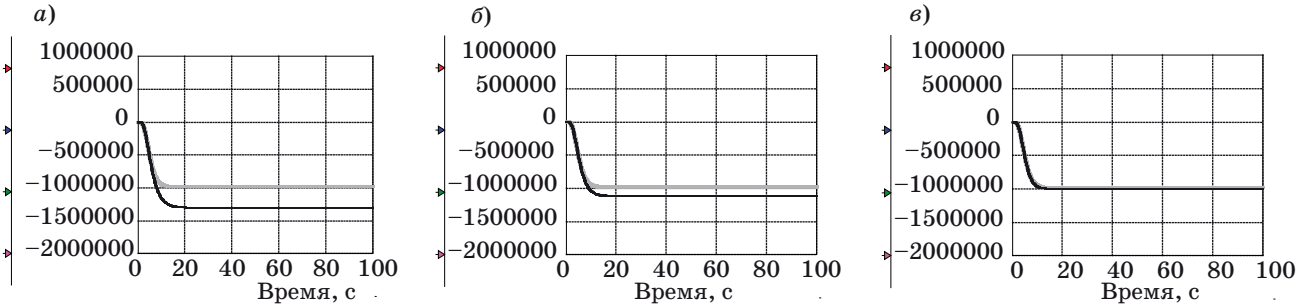
Представлены графики эталонной и исследуемой кривых для «сильной» связи 3–2 при подаче возмущения $K_{32} = 0{,}001$ без управления (рис. 5, *а*), при подключении одиночного управления (рис. 5, *б*) и полном подключении управления (рис. 5, *в*).

Моделирование показало, что подключение всех управляющих элементов при одиночной подаче возмущения в целом компенсирует малое структурное возмущение приблизительно в той же степени, как и одиночное подключение управляющего элемента, соответствующего возмущаемому узлу.

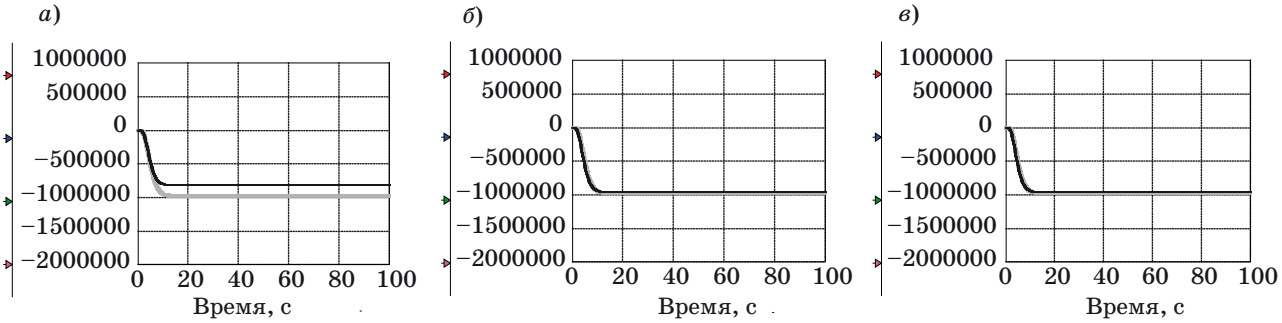
При одиночном подключении блока координации на возбуждаемый узел с одиночным подключением управления диапазоны устойчивости «слабых» связей не изменялись; как и в предыдущем случае, реакцию на воздействие проявляли четвертый, пятый, седьмой и восьмой узлы сети (помимо возмущаемого), а для «сильных» связей диапазон устойчивости резко сузился, система становилась неустойчивой при подаче даже малого возмущения. На рис. 6, *а* приведены графики кривых для той же связи 3–2, явно видно увеличение расхождения кривых по сравнению с предыдущим случаем, что говорит о сужении диапазона сходимости.

При полном подключении управления и одиночном подключении координатора (рис. 6, *б*), как и при подключении координатора на все узлы сети и одиночном подключении управления (рис. 6, *в*), диапазон устойчивости «слабых» связей не изменился, для «сильных» связей, замкнутых на первый узел, несколько расширился, для остальных — сузился.

Полное подключение управления и координации (на все узлы сети) для «слабых» связей прак-



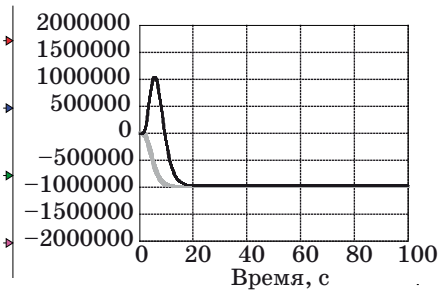
■ ***Рис. 5.*** *Эталонная и исследуемая кривые для «сильной» связи 3–2: а — без управления (установившаяся погрешность 33,3 %); б — при подключении одиночного управления; (установившаяся погрешность 14,5 %); в — полное подключение управления (установившаяся погрешность 2,14 %):* ——— *— идеальная;* ——— *— управляемая*



■ ***Рис. 6.*** *Эталонная и исследуемая кривые для «сильной» связи 3–2: а — одиночное управление и одиночная координация (установившаяся погрешность 16,67 %); б — полное управление и одиночная координация; в — одиночное управление и полная координация (установившаяся погрешность для случаев б) и в) 1,95 %):* ——— *— идеальная;* ——— *— управляемая*

тически ничего не изменило с точки зрения величины диапазона устойчивости, реакцию на возмущение помимо возмущаемого узла проявлял только восьмой узел (в предыдущих экспериментах такую реакцию проявляли еще четвертый, пятый и седьмой узлы). Для «сильных» связей диапазон устойчивости расширился, в целом реакция системы на возмущение стала слабее. На рис. 7 графики соответствуют состояниям возмущенной системы подключенным управлением и координацией для «сильной» связи 3–2 при подаче возмущения $K_{32} = 0{,}001$.

Таким образом, полное подключение управления и координации дает наилучший эффект при компенсации малых структурных возмущений.



■ ***Рис. 7.*** *Эталонная (———) и исследуемая (———) кривые для «сильной» связи 3–2, полное управление и полная координация (установившаяся погрешность 0,97 %)*

Заключение

Результаты моделирования показали, что при пошаговом изменении управляющих воздействий на отдельные линейные звенья с использованием в качестве «стабилизирующего» значения обобщенного критерия затрат, вычисляемого для каждой подсистемы на каждом шаге моделирования, за заданное время подсистемы и система в целом стремятся к «эталонным» значениям (серые кривые на графиках) выходных переменных.

Экспериментально подтверждены выводы, сделанные в работах [7, 3], об устойчивости результатов децентрализованного управления на основе градиентов локальных критериев качества.

Для исследованной двухуровневой системы управления линейным объектом подключение нижнего уровня управления в среднем на порядок расширяет диапазоны устойчивости системы к внешним возмущениям и примерно вдвое уменьшает процент расхождения идеальной и реальной траекторий системы. Координация иерархической системы позволяет:

— повысить устойчивость системы к внешним возмущениям и увеличить быстродействие системы приблизительно вдвое;

— расширить диапазон устойчивости системы к структурным возмущениям в 1,5–2 раза.

Для исследованной двухуровневой системы управления сетевым объектом выявлено разделение внутренних обратных связей на «сильные» (возмущение, подаваемое на эти связи, существенно влияет на поведение системы в целом) и «слабые». Поочередное подключение локальных управлений в среднем на порядок расширяет диапазоны устойчивости соответствующих «сильных» связей к внешним возмущениям и практически не влияет на «слабые» связи. Локальное одиночное управление тем узлом сети, на который подано возмущение, достаточно эффективно компенсирует небольшие структурные возмущения и более чем вдвое уменьшает процент расхождения идеальной и реальной траекторий системы в целом. Полное подключение нижнего уровня управления ведет к резкому сужению диапазонов устойчивости «сильных» связей (в 2–4 раза) и появлению реакции на невозбуждаемых узлах сети. Следовательно, некоординируемые локальные управления «мешают» друг другу, что и мож-

но было предположить с учетом особенностей сетевых структур.

Использование предложенного градиентного метода координации для сетевого объекта позволяет:

— повысить устойчивость системы к внешним возмущениям и свести к минимуму взаимное влияние узлов сети;

— расширить диапазон устойчивости системы к структурным возмущениям более чем в 2 раза (по сравнению с локальным управлением).

Направления дальнейших исследований:

— изучение реакции исследованных иерархической и сетевой систем на внешние возмущения в различных точках воздействия;

— поиск конструктивного алгоритма выбора оптимальных значений коэффициентов усиления k_i в цепях ввода координирующих сигналов (подбор производился вручную);

— исследование возможностей повышения быстродействия децентрализованных систем управления сетевыми структурами;

— анализ возможностей применения разработанной методики для исследования интеллектуальных динамических систем [3, 8, 9].

Работа выполнена при финансовой поддержке РФФИ (проект № 09-07-00066), Отделения нанотехнологий и информационных технологий РАН (проект 2.3 в рамках текущей Программы фундаментальных научных исследований).

Литература

1. **Месарович М., Мако Д., Такахара И.** Теория иерархических многоуровневых систем. — М.: Мир, 1973. — 344 с.
2. **Фридман А. Я.** Условия координируемости двухуровневого коллектива динамических интеллектуальных систем: Одиннадцатая национальная конф. по искусственному интеллекту с международным участием КИИ-2008, Дубна, Россия, 28 сентября — 3 октября 2008 г. — М.: ЛЕНАНД, 2008. Т. 1. С. 25–31.
3. **Фридман А. Я.** Достаточные условия координируемости локально организованной иерархии динамических систем // Искусственный интеллект. Интеллектуальные системы (ИИ-2009): Материалы X Междунар. науч.-техн. конф. / ТТИ ЮФУ. Таганрог, 2009. С. 115–117.
4. **Дьяконов В. П.** VisSim+Mathcad+MATLAB. Визуальное математическое моделирование. — М.: СОЛОН-Пресс, 2004. — 384 с. (Сер. Полное руководство пользователя).
5. **Sokolov B., Fridman A.** Integrated Situational Modelling of Industry-Business Processes for Every Stage of Their Life Cycle // Intelligent Systems (IS 2008): Proc. of 4th Intern. IEEE Conf., Sept. 6–8, 2008, Varna, Bulgaria. Vol. 1. P. 8–40.
6. **Фридман А. Я.** Ситуационный подход к моделированию промышленно-природных комплексов и управлению их структурой // Идентификация систем и задачи управления: Тр. IV Междунар. конф./ Ин-т проблем управления им. В. А. Трапезникова. М., 2005. С. 1075–1108.
7. **Стефанюк В. Л.** Локальная организация интеллектуальных систем. — М.: Физматлит, 2004. — 328 с.
8. **Виноградов А. Н., Жилякова Л. Ю., Осипов Г. С.** Динамические интеллектуальные системы. II. Моделирование целенаправленного поведения// Изв. РАН. Теория и системы управления. 2003. № 1. С. 87–94.
9. **Фридман А. Я.** Прямое планирование в динамических интеллектуальных системах // Системный анализ и информационные технологии (САИТ-2007): Тр. Второй Междунар. конф., Обнинск, Россия, 10–14 сентября 2007 г.: В 2 т. / ЛКИ. М., 2007. Т. 1. С. 73–75.

УДК 004.094

КОМБИНИРОВАНИЕ МЕХАНИЗМОВ ЗАЩИТЫ ОТ СКАНИРОВАНИЯ В КОМПЬЮТЕРНЫХ СЕТЯХ

А. А. Чечулин,
младший научный сотрудник
И. В. Котенко,
доктор техн. наук, профессор
Санкт-Петербургский институт информатики и автоматизации РАН

Предлагается подход к комбинированию различных механизмов обнаружения сетевого сканирования в компьютерных сетях, который позволяет существенно повысить эффективность обнаружения за счет уменьшения количества ложных срабатываний и повышения точности обнаружения сканирования. Дается представление об отдельных частных методиках обнаружения сканирования. Рассматриваются основные принципы их комбинирования, а также дополнительные архитектурные улучшения общей модели обнаружения сканирования. Предлагается подход к автоматической настройке параметров используемых механизмов на основе статистических данных об анализируемом трафике.

Ключевые слова *— защита информации, компьютерные сети, сетевое сканирование, обнаружение сетевых атак, комбинирование механизмов защиты.*

Введение

Актуальной задачей защиты информации в компьютерных сетях является исследование перспективных механизмов обнаружения и предотвращения сканирования сети. Сканирование сетей может использоваться как злоумышленниками для сбора информации об атакуемой компьютерной сети, так и сетевыми червями в процессе их распространения.

Представляется, что при большом количестве узлов в контролируемой компьютерной сети и многообразии работающих в ней приложений перспективным подходом к обнаружению и защите от сетевых атак является согласованное использование комплекса различных механизмов защиты [1, 2]. Это объясняется, в первую очередь, тем, что отдельные методы защиты ориентированы главным образом на определенный тип трафика и виды атак, а комбинированные механизмы позволяют объединить преимущества отдельных методов и нивелировать их недостатки.

В статье предлагается и исследуется подход к комбинированию различных механизмов обнаружения сетевого сканирования в компьютерных сетях, который позволяет значительно повысить эффективность обнаружения. Хотя в существующих работах (например, в [3–9]) задача разработки отдельных частных механизмов обнаружения сканирования в достаточной степени исследована, проблема повышения эффективности этих механизмов и возможности их автоматической настройки в соответствии с текущей сетевой обстановкой (используемыми в сети приложениями, параметрами трафика и т. п.), в том числе за счет комбинирования, остается не решенной.

Решение поставленной проблемы в статье представлено на основе последовательного выполнения следующих задач: определения нескольких эффективных механизмов защиты; выделения наиболее важных параметров трафика и классификации трафика по ним; определения показателей эффективности обнаружения; разработки общего подхода к комбинированию механизмов защиты, в том числе разработки и обучения алгоритма подбора оптимальных параметров для механизмов на каждом из выделенных классов трафика и разработки и обучения алгоритма комбинирования механизмов защиты.

Механизмы обнаружения сканирования

За основу предлагаемого комбинированного подхода к обнаружению сетевого сканирования в компьютерных сетях было принято использование нескольких семейств механизмов, базирующихся

на следующих методиках: методике «дросселирования/регулирования вирусов» (*Virus Throttling*) и ее модификации (VT-S и VT-C) [4, 5]; методиках, основанных на анализе неудачных соединений (*Failed Connection* — FC) [6]; методиках, использующих метод «порогового случайного прохождения» (*Threshold Random Walk* — TRW) [7, 8]; методиках ограничения интенсивности соединений на основе кредитов доверия (*Credit Based Rate Limiting* — CB) [9]. Представим основные элементы реализации нескольких из указанных механизмов.

Методика «дросселирования/регулирования вирусов» для реализации на коммутаторах (VT-S) основывается на следующей модели обработки сетевого трафика. Для каждого узла существует список длиной N для отметки значений хэш-функции на последние адреса, к которым приходили запросы на соединения. Хэш-функция принимает значения от 0 до $N - 1$, при получении значения k на k-е место в списке записывается 1. Запрос на соединение идентифицируется по пакету TCP-SYN. Задается минимальный интервал между поступлениями запросов на соединение к новым адресам (rate threshold). Если запрос поступил спустя время, меньшее, чем заданный порог, адрес получателя хэшируется и происходит запись в таблицу хэшей для этого отправителя. В противном случае список отметок хэшей очищается и добавляется только информация по текущему запросу. Если таблица хэшей полна, то все запросы от отправителя считаются вредоносными на время T. Используется протокол TCP, анализируются пакеты TCP SYN. В пакетах контролируются поля «source IP» и «destination IP».

Предлагаемая модификация методики Virus Throttling на основе метода CUSUM (VT-C) [3] заключается в следующем. Для каждого узла существует счетчик C. $CMax$ — это порог количества запросов к новым адресам подряд. Счетчик увеличивается, если прошло менее $RMin$ секунд с момента прихода последнего запроса с незарегистрированным адресом получателя. Запрос на соединение идентифицируется по пакету TCP-SYN. При превышении счетчиком C заданного порога запросы на соединение от данного узла считаются вредоносными. Так же как и для методики Virus Throttling, используется протокол TCP и анализируются пакеты TCP SYN. В пакетах контролируются поля «source IP» и «destination IP».

В *методиках* FC, основанных на анализе неудачных соединений, используется и хранится следующая информация: адрес узла, инициировавшего соединение, частота ошибок, время соединения и счетчик ошибочных соединений. При превышении порога частоты добавления новых записей в хэш-таблицу создается запись для индивидуального узла. В дальнейшем при увеличении значения счетчика ошибочных соединений на фиксированную величину (например, 100) для этой записи обновляется значение частоты ошибок f (failure rate). При превышении величины failure rate заданного порога запросы на соединение от данного узла считаются вредоносными. При реализации данных методик используется протокол TCP, анализируются пакеты TCP RESET. В пакетах контролируются поля «source IP» и флаг TCP RST.

Методика TRW базируется на следующей модели вычислений. Для анализа узла, демонстрирующего большую сетевую активность, на предмет возможности проведения сканирования используется метод последовательного тестирования гипотез (*Sequential Hypothesis Testing*).

Пусть H_1 — гипотеза, что узел r демонстрирует повышенную сетевую активность (проводит сканирование), H_0 — гипотеза, что узел не демонстрирует повышенной сетевой активности (не проводит сканирование), а Y_i — переменная, характеризующая результат попытки соединения с i-м узлом, к которому направлен запрос соединения. Эта переменная может иметь следующие значения: 0, если соединение установлено; 1, если соединение не установлено.

Предполагается, что условия наступления гипотез H_i — случайные величины $Y_i \,|\, h_i = 1, 2, \ldots$ — независимы и распределены равномерно. В этом случае можно выразить распределение бернуллиевой случайной величины Y_i так:

$$P_r = [Y_i = 0 \,|\, H_0] = \theta_0,\ P_r = [Y_i = 1 \,|\, H_0] = 1 - \theta_0,$$
$$P_r = [Y_i = 0 \,|\, H_1] = \theta_1,\ P_r = [Y_i = 1 \,|\, H_1] = 1 - \theta_1.$$

Предположение, что попытка установления соединения от неинфицированного узла имеет большую вероятность успешного исхода, чем от инфицированного, подразумевает выполнение условия $\theta_0 > \theta_1$.

Учитывая эти две гипотезы, при принятии решения о вредоносности/безопасности узла возможны следующие четыре варианта:

1) «обнаружение сканирования» — выбирается гипотеза H_1 при реальном сканировании с узла;

2) «пропуск атаки (false negative)» — выбирается гипотеза H_0 при реальном сканировании с узла;

3) «отсутствие сканирования» — выбирается гипотеза H_0 при реальном отсутствии сканирования с узла;

4) «ложное срабатывание (false positive)» — выбирается гипотеза H_1 при реальном отсутствии сканирования с узла.

Введем обозначения P_D — вероятность обнаружения сканирования с узла и P_F — вероятность

ложного срабатывания с целью определить требования к выполнению методики. Для максимальной эффективности выполнения методики предполагается, что верны следующие условия: $P_D \geq \beta$, $P_F \leq \alpha$, $\alpha = 0{,}01$, $\beta = 0{,}99$.

При наступлении анализируемого события вычисляется следующее вероятностное отношение:

$$\Lambda(\mathbf{Y}) \equiv \frac{\Pr[\mathbf{Y} \mid H_1]}{\Pr[\mathbf{Y} \mid H_0]} = \prod_{i=1}^{n} \frac{\Pr[Y_i \mid H_1]}{\Pr[Y_i \mid H_0]},$$

где $\mathbf{Y}$ — вектор наблюдаемых событий; $P_r[Y_i \mid H_i]$ — условная вероятность наступления события, при котором последовательность $\mathbf{Y}$ полностью соответствует гипотезе H_i.

Затем выполняются сравнения вероятностного отношения Λ с верхним (η_1) и нижним (η_0) пороговыми значениями и по результатам сравнения принимается решение о наличии/отсутствии сканирования с узла: $\Lambda(\mathbf{Y}) \leq \eta_0 \Rightarrow H_0$, $\Lambda(\mathbf{Y}) \geq \eta_1 \Rightarrow H_1$. В случае если $\eta_0 < \Lambda(\mathbf{Y}) < \eta_1$, продолжается ожидание дополнительных событий для более точной идентификации. Пороговые значения выбираются следующим образом:

$$\eta_1 = \frac{\beta}{\alpha} \text{ и } \eta_0 = \frac{1-\beta}{1-\alpha}.$$

Выбор параметров сетевого трафика

Для классификации трафика в данной статье выделено около 30 параметров. В качестве основных были выбраны следующие: интенсивность соединений; процент запросов на установление соединения (TCP SYN пакеты) от общего количества пакетов; продолжительность соединений; интервалы между приходами пакетов в соединении; среднее количество пакетов на один хост-источник; отношение количества запросов на установление соединений (TCP SYN) с количеством подтверждений на установление соединений (TCP SYN-ACK) и др.

Задача обучения, заключающаяся в поиске оптимальных конфигураций используемых отдельных механизмов защиты на основе параметров трафика, объективно осложнена высокой корреляцией этих параметров в реальных сетях.

Выбранные параметры трафика (средняя частота приходящих пакетов, процент TCP SYN запросов и др.) представляют собой некоторые статистические данные, собранные за определенный период времени.

Существует несколько подходов к вычислению таких параметров. Наиболее простой подход — это хранение в памяти всех пакетов за анализируемый период времени и вычисление статистических параметров по этой истории. Основным недостатком его является большая ресурсоемкость, особенно для высокоскоростных сетей.

Другой вариант — хранение только самих значений статистических параметров. Например, значение средней частоты трафика за период можно вычислить, имея количество пришедших пакетов с начала периода. Этот путь значительно менее ресурсоемкий, однако по истечении разумно небольшого периода статистические значения приходится обнулять, так как средние значения показателя (например, за неделю) уже использовать бессмысленно. После обнуления параметров и до их последующего расчета использование методов комбинирования невозможно.

Для преодоления этой ситуации предлагается считать статистику параметров трафика сразу в нескольких временных «окнах», образуемых со сдвигом по времени. При вычислениях используются параметры трафика самого «старого» окна. Количество окон может задаваться произвольно при настройке системы.

Показатели эффективности механизмов обнаружения

Показатели эффективности механизмов обнаружения хостов со сканирующим трафиком основаны на бинарной классификации хостов вида «содержит/не содержит» вредоносный трафик. Данные показатели основываются на следующей матрице классификации:

	Реально содержит	Реально не содержит
Распознан как вредоносный	*a* (true positive)	*b* (false positive)
Распознан как не вредоносный	*c* (false negative)	*d* (true negative)

Здесь *a* — количество хостов, трафик которых содержит сканирующий трафик и правильно распознанных системой; *b* — количество хостов, трафик которых не содержит сканирующий трафик, но распознанных системой как вредоносные; *c* — количество хостов, трафик которых содержит сканирующий трафик, но распознанных системой как безопасные; *d* — количество прочих хостов.

В проводимых авторами исследованиях по анализу механизмов обнаружения сканирования, наряду с показателями ошибок первого и второго рода — степенью ложных срабатываний (false positive) и степенью пропусков атак (false negative), используются также следующие интегрированные показатели, вычисляемые на основе показателей ошибок первого и второго рода: полнота, точность, аккуратность, *F*-мера и ошибка.

Полнота r вычисляется как отношение правильно распознанных вредоносных хостов к обще-

му количеству вредоносных хостов. Этот параметр характеризует способность системы распознавать вредоносные хосты, но не учитывает количество неправильно распознанных безопасных хостов.

Точность p определяется как отношение правильно распознанных вредоносных хостов к общему количеству хостов, распознанных как вредоносные. Точность характеризует способность системы распознавать только реально вредоносные хосты.

Аккуратность вычисляется как отношение правильно принятых системой решений к общему числу решений.

F-мера используется как единая метрика, объединяющая метрики полноты и точности; вычисляется по формуле $F = 2pr/(p + r)$.

Ошибка определяется как отношение неправильно принятых системой решений к общему числу решений.

В статье для определения эффективности механизмов обнаружения используются показатели false positive, false negative, *F*-меры и аккуратности.

Сущность подхода к комбинированию механизмов обнаружения

Комбинирование механизмов обнаружения и реагирования основывается на использовании в процессе работы тех механизмов, которые проявили себя наилучшим образом при обучении на трафике, максимально близком по всем параметрам к тому, на котором происходит реальное обнаружение сканирования.

Для решения задачи комбинирования применялись следующие методы интеграции конечных результатов отдельных механизмов:

1) выбор наилучшего (оптимального) механизма и отключение остальных;

2) простого большинства голосов;

3) взвешенного большинства голосов;

4) комбинирование с помощью методов Data Mining.

Все предложенные методы предполагают первоначальное тестирование каждого метода защиты и нахождение лучшего метода (с заданными параметрами) при определенных характеристиках трафика. Предполагается, что данные методы позволяют использовать наиболее сильные стороны каждого механизма для каждого класса трафика.

Сложностью реализации *первого метода* является то, что не всегда можно осуществить полное упорядочение механизмов защиты по их степени применимости, в первую очередь, в силу разнообразия возможного трафика и его характеристик.

Второй метод использует данные от всех механизмов защиты, но при этом предполагает, что механизмы работают с одинаковой степенью точности и, как и первый метод, не учитывает степень применимости механизмов защиты для различных ситуаций.

Представим более детально предлагаемый вариант реализации *третьего метода*, основанного на вычислении весовых коэффициентов для каждого механизма.

Весовые коэффициенты определяются на основе параметров трафика, которые должны быть выбраны заранее. Функция $W(d, x_1, ..., x_n)$ вычисления весовых коэффициентов принимает в качестве аргументов тип (класс) d механизма защиты и точку декартова n-мерного гиперпространства, заданного параметрами $(x_1, ..., x_n)$. Размерность гиперпространства определяется количеством выбранных параметров трафика.

По результатам предварительного тестирования (обучения) для каждого параметра строится зависимость значения *F*-меры при использовании различных механизмов защиты от значения определенного параметра.

Определим функцию $F(x)$ зависимости значения *F*-меры от параметра трафика x следующим образом: на основе данных, полученных на этапе обучения, строится интерполяционный полином Лагранжа, который позволяет получить предположительные значения *F*-меры для любого значения параметра x.

Обозначим через $F_i(d, x)$ функцию зависимости значения *F*-меры от i-го параметра для механизма защиты d. Тогда функцию вычисления весовых коэффициентов определим следующим образом:

$$W(d, x_1, \ldots, x_n) = \frac{\sum_{i=1}^{n} F_i(d, x)}{n}.$$

Значение функции W тем меньше, чем меньше значение *F*-меры для данного механизма при данных характеристиках трафика.

При вычислении данной функции предполагается, что все параметры трафика являются равнозначными. Можно также снабдить параметры трафика весами согласно их важности для принятия решения. Тогда слагаемые суммы в числителе следует умножить на соответствующий весовой коэффициент.

Пусть функция $HF(f, d)$ принятия решения о вредоносности хоста, которая получает значение после получения пакета f механизмом защиты d, равна –1, если хост признан вредоносным, и +1, если хост признан безопасным.

Тогда функция голосования для m механизмов защиты при принятии решения о вредоносности хоста после получения пакета f при характеристиках трафика $x_1, ..., x_n$ выглядит следующим образом:

$$V(f, x_1, \ldots, x_n) = \sum_{j=1}^{m} (W(d_j, x_1, \ldots, x_n) HF(f, d_j)).$$

Если значение $V(f, x_1, ..., x_n)$ отрицательно, то хост признается вредоносным, в противном случае — безопасным.

Четвертый метод комбинирования использует данные от всех механизмов защиты, и эти данные вместе с основными параметрами трафика передаются в обученный заранее классификатор (например, Naïve Bayes [10]). В результате на основе входных данных и предварительного обучения классификатор принимает решение о вредоносности анализируемого трафика.

Эти же методы комбинирования применимы и к автоматическому выбору оптимальных параметров механизмов обнаружения сканирования. Основная идея данного подхода заключается в том, что целесообразно также распространить комбинирование на разные конфигурации отдельных механизмов, меняя параметры настройки механизма в зависимости от текущей ситуации в сети.

Основные требования к разрабатываемым комбинированным механизмам защиты в целом соответствуют функциональным требованиям разработанного проактивного подхода к обнаружению сканирования в целом [2], а именно адекватность и оперативность обнаружения, эффективность использования системных ресурсов, автоматическое выполнение, возможность обнаружения различных видов сетевых червей.

Среда для проведения экспериментов

Для моделирования и оценки методов комбинирования механизмов защиты и отдельных механизмов защиты, а также выбора для них оптимальных параметров разработана автоматизированная методика и программные средства исследования (моделирования и анализа) механизмов защиты.

Суть практической оценки методов комбинирования и механизмов защиты сводится к проведению комплекса экспериментов на основе использования программной среды для различных значений входных параметров и измерению показателей эффективности исследуемых механизмов защиты.

При реализации программных средств моделирования и анализа отдельных механизмов, методов комбинирования механизмов и методики выбора оптимальных параметров для механизмов использовалась *архитектура*, включающая следующие компоненты:

- *модели источников трафика,* предназначенные для предоставления сетевого трафика исследуемым механизмам защиты, включая как модели трафика сканирования, так и модели обычного сетевого трафика;
- *модели предобработки и синхронизации источников трафика,* служащие для приведения трафика из формата источников в формат, удобный для анализа исследуемыми механизмами защиты. Этот модуль также предназначен для синхронизации трафика при одновременном использовании нескольких источников трафика;
- *модели механизмов обнаружения сканирования*. Входными параметрами каждого механизма обнаружения являются те поля сетевого пакета, полученного от источника трафика, которые им обрабатываются. В качестве управляющих параметров вводятся различные внутренние параметры каждого механизма, которые влияют на его эффективность.

Реализованные программные средства моделирования и анализа механизмов защиты позволяют оценивать следующие основные показатели эффективности методов комбинирования и выбора оптимальных параметров:

- долю блокированного и (или) задержанного легитимного трафика (степень ложных срабатываний, false positives);
- долю пропущенного злонамеренного трафика (степень пропусков атак, false negatives);
- интегрированные показатели (полноту, точность, аккуратность, *F*-меру и ошибку);
- время реакции на атаку.

При проведении экспериментов использовался комбинированный способ моделирования трафика, заключающийся в применении в качестве входных данных различных записей реального трафика с дополнением их необходимым для исследования трафиком. В данном случае необходимый для исследования трафик — это трафик сканирования и трафик «быстрых» приложений, таких как P2P или NetBIOS/NS. Исследования проводились как на записях реального трафика, так и на сгенерированном трафике с подключением различных модулей генерации моделей трафика.

Для сравнения механизмов использовались трафики двух типов:

- трафик сети уровня предприятия с большим набором различных приложений (в том числе сканеров уязвимостей) без преобладания трафика какого-либо из них;
- трафик локальной сети с преобладанием приложений P2P.

Выбор таких типов трафика обоснован тем, что механизмы обнаружения и реагирования против сетевого сканирования должны их точно обнаруживать как в трафике легитимных приложений, так и в трафике приложений, создающих большое количество соединений с различными узлами. Последнее особенно важно, так как трафик приложений, создающих большое количе-

ство соединений с различными узлами, похож на трафик, появляющийся при сканировании.

Результаты экспериментов

Для проведения экспериментов указанные виды трафика были смешаны с трафиком сканирования. Использовались следующие параметры сканирования: скорость отправки запросов на соединение — 50 запросов/с, вероятность успешного соединения — 30%, вероятность получения TCP-RST пакета — 30%, выбор IP-адресов для сканирования — случайный.

Приведем сначала результаты тестирования отдельных механизмов обнаружения (VT-S, VT-C, FC, TRW и CB) для различных видов трафика.

Средние значения процента ложных срабатываний (FP) и процента пропуска атак (FN) для отдельных механизмов на обычном и P2P-трафике без трафика сканирования представлены в табл. 1.

Средние значения процента ложных срабатываний (FP), процента пропуска атак (FN), объема памяти (V) и аккуратности (A) для отдельных механизмов на обычном и P2P-трафике, смешанном с трафиком сканирования, показаны в табл. 2.

Отметим, что во многих случаях для трафика P2P предложенный механизм Virus throttling на основе метода CUSUM (VT-C) превзошел используемый в настоящее время на сетевых коммутаторах механизм VT-S. Этот факт свидетельствует о целесообразности его реализации на коммутаторах и использовании его вместо VT-S или совместно с VT-S в сетях с трафиком P2P.

В настоящее время авторы продолжают проводить эксперименты для различных реализаций методов комбинирования механизмов защиты.

Проведенные эксперименты показали, что использование методов комбинирования, основанных на выборе наилучшего (оптимального) механизма, применении взвешенного большинства голосов и методов Data Mining для различных типов трафика приводит в большинстве случаев к улучшению показателей эффективности.

Эти эксперименты были проведены на смеси обычного трафика, трафика, содержащего соединения P2P, и трафика сетевого червя таким образом, чтобы основные параметры хостов заметно менялись. Это осуществлялось, например, путем добавления в трафик хостов с некоторого момента времени трафика P2P или, наоборот, удаления трафика P2P с хоста, на котором до этого времени этот трафик был.

При использовании метода комбинирования, основанного на выборе оптимального механизма, для обычного трафика был автоматически выбран метод VT-C, а для трафика, содержащего P2P-трафик — метод CB. К сожалению, реализованный метод комбинирования не учитывал изменений в параметрах трафика, и при появлении в обычном трафике трафика P2P и, наоборот, в случае удаления трафика P2P комбинированный механизм начинал показывать результаты, заметно худшие даже в сравнении с некоторыми отдельными методами обнаружения сканирования.

Метод комбинирования на базе простого большинства голосов лучше реагировал на изменение параметров трафика, чем отдельные методы обнаружения, но результаты в каждом конкретном случае оказались хуже работы отдельных методов, оптимальных для данных параметров трафика.

■ ***Таблица 1.*** *Результаты работы механизмов защиты на трафиках без сканирования*

Механизм защиты	Обычный трафик		P2P-трафик	
	FP	FN	FP	FN
VT-S	0,022600	0	0,089913	0
VT-C	0,023400	0	0,061528	0
FC	0,005950	0	0,002946	0
TRW	0,004300	0	0,002311	0
CB	0,021800	0	0,080051	0

■ ***Таблица 2.*** *Результаты работы механизмов защиты на трафиках, смешанных с трафиком сканирования*

Вид трафика	Механизм защиты	FP	FN	V	A
Обычный трафик	VT-S	0,023300	0,000663	23496	0,998929
	VT-C	0,024100	0,000530	36224	0,998983
	FC	0,009680	0,054291	48572	0,972348
	TRW	0,004980	0,002416	28912	0,998216
	CB	0,025600	0,000344	9608	0,987114
P2P-трафик	VT-S	0,301541	0,0178923	34820	0,96833
	VT-C	0,101948	0,0317885	105272	0,964804
	FC	0,103761	0,013333	74108	0,98227
	TRW	0,0599373	0,0972275	127656	0,905997
	CB	0,200021	0,001258	22632	0,993393

При использовании метода комбинирования, основанного на взвешенном большинстве голосов, в качестве отслеживаемых параметров были выбраны интенсивность запросов на установление соединений, процент запросов на установление соединения (пакеты TCP SYN) от общего количества пакетов, продолжительность соединений, среднее количество пакетов на один хост-источник. В результате проведения экспериментов были получены следующие средние значения: процента ложных срабатываний FP = 0,009421 и процента пропуска атак FN = 0,008596.

Наилучшие результаты были показаны при применении для комбинирования методов Data Mining (в экспериментах использовался наивный байесовский подход). В результате проведенных экспериментов получены средние значения процента ложных срабатываний FP = 0,003529 и процента пропуска атак FN = 0,001286.

На процент ложных срабатываний влияет то, что после обнаружения сканирования с хоста хост блокируется, при этом блокируется и его нормальный (не вредоносный) трафик.

Заключение

В данной статье предложен подход к комбинированию механизмов обнаружения сканирования и выбору оптимальных параметров для этих механизмов. Представлено разработанное программное средство для проведения экспериментов и определения эффективности средств защиты на записях реального трафика.

По результатам экспериментов проведена оценка отдельных и комбинированных механизмов обнаружения сканирования. По полученным данным можно судить о том, что использование предложенных методов комбинирования, в частности метода комбинирования, основанного на взвешенном большинстве голосов, и методов Data Mining, а также настройка параметров для отдельных механизмов защиты в зависимости от статистических показателей трафика позволяет существенно улучшить эффективность работы этих механизмов.

В дальнейшей работе планируется проведение большого количества экспериментов для анализа эффективности методов комбинирования для различных смесей исследуемого сетевого трафика, разработка новых и совершенствование существующих отдельных механизмов обнаружения, исследование других схем кооперации отдельных механизмов, развитие разработанного программного средства в целях создания среды моделирования для механизмов защиты от сканирования.

Работа выполняется при финансовой поддержке РФФИ (проект № 10-01-00826-а), программы фундаментальных исследований ОНИТ РАН (проект № 3.2) и при частичной финансовой поддержке, осуществляемой в рамках проектов Евросоюза SecFutur и MASSIF.

Литература

1. **Чечулин А. А.** Обнаружение и противодействие сетевым атакам на основе комбинированных механизмов анализа трафика // Информационная безопасность регионов России (ИБРР-2009): Материалы VI Санкт-Петербургской межрегион. конф., 28–30 октября 2009 г. / СПОИСУ. СПб., 2009. С. 143–144.
2. **Котенко И. В., Воронцов В. В., Чечулин А. А., Уланов А. В.** Проактивные механизмы защиты от сетевых червей: подход, реализация и результаты экспериментов // Информационные технологии. 2009. № 1. С. 37–42.
3. **Чечулин А. А., Котенко И. В.** Исследование механизмов защиты от сетевых червей на основе методик Virus Throttling // Защита информации. Инсайд. 2008. № 3. С. 68–73.
4. **Williamson M.** Throttling Viruses: Restricting propagation to defeat malicious mobile code // Proc. of ACSAC Security Conf., Las Vegas, Nevada. 2002. P. 61–68.
5. **Twycross J., Williamson M.** Implementing and testing a virus throttle // Proc. 12th USENIX Security Symp., 2003. http://www.hpl.hp.com/techreports/2003/HPL-2003-103.html (дата обращения 09.10.2010).
6. **Chen S., Tang Y.** Slowing Down Internet Worms // 24th Intern. Conf. on Distributed Computing Systems (ICDCS'04), Tokyo, Japan, Mar. 2004. P. 312–319.
7. **Jung J., Paxson V., Berger A. W., Balakrishnan H.** Fast portscan detection using sequential hypothesis testing // Proc. of the 2004 IEEE Symp. on Security and Privacy, Oakland, California, USA, May 9–12, 2004 / IEEE Computer Society, 2004. P. 211–225.
8. **Weaver N., Staniford S., Paxson V.** Very fast containment of scanning worms // Proc. of the 13th USENIX Security Symp., Aug. 9–13, 2004. http://www.icsi.berkeley.edu/~nweaver/containment/containment.pdf (дата обращения 09.10.2010).
9. **Schechter S., Jung J., Berger A. W.** Fast Detection of Scanning Worm Infections // Proc. of the Seventh Intern. Symp. on Recent Advances in Intrusion Detection, French Riviera, France, Sept. 2004. P. 59–81.
10. **Zuev D., Moore A. W.** Traffic Classification using a Statistical Approach: Proc. of the 2005 ACM SIGMETRICS Intern. Conf. on Measurement and Modeling of Computer Systems, Banff, Alberta, Canada, June 06–10, 2005//Lecture Notes in Computer Science. Springer, 2005. Vol. 3431. P. 321–324.

УДК 621.391

СИСТЕМА МНОЖЕСТВЕННОГО ДОСТУПА, ИСПОЛЬЗУЮЩАЯ НЕКОГЕРЕНТНЫЙ ПОРОГОВЫЙ ПРИЕМ, ЧАСТОТНО-ПОЗИЦИОННОЕ КОДИРОВАНИЕ И ДИНАМИЧЕСКИ ВЫДЕЛЯЕМЫЙ ДИАПАЗОН ЧАСТОТ, В УСЛОВИЯХ ПОДАВЛЕНИЯ ПОЛЕЗНОГО СИГНАЛА

Д. С. Осипов,
канд. техн. наук, старший научный сотрудник
Институт проблем передачи информации им. А. А. Харкевича РАН

Рассматривается модель системы множественного доступа, использующей динамически выделяемые поддиапазоны ортогональных частот, технологию частотно-позиционного кодирования и некогерентный пороговый прием в условиях подавления сигналов, передаваемых в системе. Для подавления используются сигналы, по форме и ширине занимаемой полосы аналогичные полезным. В работе проводится сравнение различных вариантов использования такой стратегии подавления с точки зрения их влияния на максимально возможную скорость надежной передачи информации рассматриваемым пользователем.

Ключевые слова *— система множественного доступа, динамически выделяемые частотные поддиапазоны, некогерентный пороговый прием, подавление.*

Введение

Одним из важнейших требований, предъявляемых к современным системам связи, является обеспечение защиты передаваемых данных от намеренного подавления. Это требование особенно актуально для систем множественного доступа, использующих беспроводные каналы связи, так как интенсивное развитие систем такого типа и непрерывный рост числа приложений, использующих принципы разделения пользователей в системах множественного доступа, обуславливают увеличение риска потери данных вследствие применения технологий подавления полезного сигнала и возможный урон от такого рода враждебной активности. Технология псевдослучайно переключающихся радиочастот (ППРЧ) в течение долгого времени рассматривалась в качестве наиболее эффективного метода противодействия технологиям подавления полезного сигнала. Однако развитие методов подавления привело к созданию новых технологий, специально предназначенных для подавления полезного радиосигнала в системах связи, использующих ППРЧ. В связи с этим в работе [1] была предложена модификация системы ППРЧ, использующая динамическое выделение частотных поддиапазонов и пороговый прием. Благодаря этим особенностям система такого типа оказывается намного лучше защищенной от существующих методов подавления с использованием узкополосных помех (по крайней мере, в случае, когда используется некогерентный прием). Вместе с тем исследование влияния различных стратегий подавления с использованием узкополосных помех на качество работы системы описанного типа до последнего времени не проводилось. Настоящая работа является первой попыткой восполнить этот пробел.

Система множественного доступа, использующая псевдослучайное переключение радиочастот

Для того чтобы пояснить особенности решаемой нами задачи, опишем более детально проблему множественного доступа в той ее форме, которая рассмотрена в статье. Приведем ситуацию, в которой некоторый пользователь (в дальнейшем именуемый «рассматриваемым») передает данные на базовую станцию. Предположим, что од-

новременно с рассматриваемым пользователем передачу могут вести еще K пользователей (таких пользователей станем называть «мешающими», а всех пользователей, передающих данные в такой системе, «активными»), которые руководствуются при передаче теми же принципами, что и рассматриваемый пользователь, и при этом передают информацию асинхронно и некоординированно.

Рассмотрим систему множественного доступа, использующую для разделения пользователей технологию ППРЧ. Предположим, что полоса частот, предоставленная пользователям, разбита на неперекрывающиеся частотные поддиапазоны. Каждый пользователь оснащен генератором, псевдослучайно выбирающим номер поддиапазона, в котором будет вестись передача. Собственно передача ведется с использованием заранее выбранного метода модуляции (например, частотной модуляции). Переключение между выбранными поддиапазонами традиционно именуется прыжком. Обозначим длительность промежутка между двумя прыжками T_h.

Предполагается, что приемник оснащен псевдослучайными генераторами, каждый из которых засинхронизирован с псевдослучайным генератором одного из пользователей. Последнее означает, что в момент приема сигнала, соответствующего сигналу, переданному рассматриваемым пользователем, генератор выдает номер поддиапазона такой же, как и тот, что был выдан генератором рассматриваемого пользователя (заметим, что условие синхронизации генераторов означает наличие кадровой синхронизации. Иными словами, несмотря на то, что пользователи передают информацию асинхронно, каждая пара «передатчик—приемник» должна быть засинхронизирована в вышеуказанном смысле), что позволяет приемнику определить частотный поддиапазон, который был использован рассматриваемым пользователем для передачи, а затем демодулировать принятый сигнал и принять решение о переданном символе.

Традиционно при рассмотрении систем, построенных в соответствии с вышеописанным принципом, предполагалось, что псевдослучайные номера, вырабатываемые генераторами, неизвестны никому, кроме пары «передатчик—приемник». Тем самым предполагалось, что пользователь, целью которого является подавление полезного сигнала (в дальнейшем будем называть таких пользователей «враждебными»), не может определить, какие именно диапазоны используются для передачи, и, соответственно, сформировать сигнал подавления. Однако в настоящее время появились новые типы генераторов намеренных помех, позволяющие определить поддиапазоны, используемые для передачи (этот процесс занимает сравнительно небольшую долю от величины T_h), и сформировать помеху, не позволяющую корректно принять переданный символ (такая стратегия подавления оказывается тем более эффективной, чем выше порядок модуляции, который использует рассматриваемый пользователь).

В работе [1] была предложена модификация вышеописанной системы множественного доступа, использующая динамически выделяемые частотные диапазоны. Опишем более подробно модифицированную систему такого типа, использующую пороговый прием (пример системы такого типа, использующей другой способ приема, можно найти в работе [2]).

Система множественного доступа, использующая динамически выделяемые частотные диапазоны, ППРЧ и пороговый прием

Рассмотрим ситуацию, в которой $\tilde{K} = K + 1$ пользователей ведут передачу данных на базовую станцию независимо, асинхронно и некоординированно, руководствуясь одними и теми же принципами. Пусть используемая полоса частот разбита на Q непересекающихся частотных подканалов (проще и эффективнее всего реализовать это условие, применяя технологию ортогонального мультиплексирования частот OFDM, поэтому в дальнейшем мы будем рассматривать систему, в которой используется именно эта технология). В системе и передатчик, и приемник оснащены засинхронизированными генераторами, которые псевдослучайно выбирают (без повторений) q из Q номеров подканалов. Каждый пользователь передает q-ичные символы, и каждому из них поставлен в соответствие один из выбранных генератором подканалов. При передаче i-го символа пользователь передает сигнал в подканале, поставленном в соответствие этому символу.

Используя сигналы, принятые из подканалов выбранных генератором номеров подканалов, приемник вычисляет для каждого из них некоторую статистику и сравнивает каждое из вычисленных значений с некоторым порогом. В случае если превышение порога регистрируется в одном из подканалов, то принимается символ, поставленный в соответствие этому подканалу. В противном случае принимается решение о стирании. Если принятый символ отличается от переданного, говорят об ошибке.

Таким образом, описанная система может рассматриваться как модифицированный вариант системы с ППРЧ, использующей частотное манипулирование и пороговый прием. Отличие состоит в том, что в этом варианте поддиапазоны фик-

сированы, в то время как в описанной нами системе частотные поддиапазоны (представляющие собой наборы непересекающихся подканалов) выделяются динамически. В первом случае пользователь, получивший доступ к информации о характере распределения подканалов (например, вошедший в систему как легальный пользователь), может, анализируя спектр, восстановить номера поддиапазонов, которые уже используются, и, посылая сигналы в других подканалах тех же диапазонов, генерировать стирания в последовательностях, принимаемых другими пользователями. В результате пользователь, обладая достаточными энергетическими ресурсами, может сделать надежную передачу данных легальными пользователями (или, по крайней мере, значительной их части) невозможной. В системе с динамическим выделением диапазонов применение такой стратегии подавления невозможно, так как используемый для передачи каждого следующего символа поддиапазон известен лишь паре «передатчик—приемник» и выбирается вновь для передачи каждого следующего символа.

Рассмотрим систему описываемого типа, использующую прием по мощности. Такая система существенно лучше защищена от подавления сигналами, передаваемыми непосредственно в подканалах, используемых рассматриваемыми пользователями, по сравнению с системой, использующей другой вид приема или тип модуляции, так как в среднем энергия принятого сигнала возрастает. Именно эта техника подавления будет рассмотрена ниже. Для того чтобы охарактеризовать функционирование системы рассматриваемого типа в условиях подавления с использованием сосредоточенной помехи, заметим, что каждый из множества «восходящих» каналов в описанной нами системе представляет собой по сути q-ичный дискретный канал без памяти со стираниями, который в свою очередь может быть представлен как композиция симметричного q-ичного дискретного канала без памяти C1 и стирающего канала C2 (под стирающим каналом здесь подразумевается канал, в котором каждый символ переходит или сам в себя, или в стирание). Аналитическое выражение для пропускной способности канала вышеуказанного типа, характеризующегося вероятностью ошибки p_e и вероятностью стирания p_x, было получено в явном виде [3]:

$$C(p_e, p_x) = [\log_2 q + (1-\tilde{p}_e)\log_2(1-\tilde{p}_e) + \\ + \tilde{p}_e \log_2 \tilde{p}_e - \tilde{p}_e \log_2(q-1)](1-p_x),$$

где $\tilde{p}_e = \dfrac{p_e}{1-p_x}$ — вероятность ошибки в канале C1.

Следует отметить, что пропускная способность такого канала имеет смысл максимальной скорости, с которой пользователь, применяя описанный выше метод передачи, может «надежно» (здесь этот термин используется в том же смысле, что и в работе [4]) вести передачу данных по каналу, характеризуемому вероятностью ошибки p_e и вероятностью стирания p_x. Последние в действительности являются функциями как параметров, которые фиксируются на этапе проектирования (общее число подканалов; число подканалов, предоставляемых каждому из пользователей), и которыми не может управлять ни пользователь, ни проектировщик (отношение сигнал/шум, количество мешающих пользователей), так и такими параметрами, как величины порогов, которые должны выбираться. Здесь и далее будем рассматривать систему с идеальным контролем мощности, что в частности означает, что все пороги можно выбрать одинаковыми и равными Tr. Обозначим максимально возможную скорость передачи, о которой говорилось выше, как R:

$$R(Q, q, SNR, K, Tr) = C(p_e(Q, q, SNR, K, Tr), \\ p_x(Q, q, SNR, K, Tr)).$$

При условии, что все прочие параметры зафиксированы, величину порога разумно выбрать таким образом, чтобы R максимизировалась. Полученное максимальное значение

$$R_m = R_m(Q, q, SNR, K) = \max_{Tr} R_m(Q, q, SNR, K, Tr).$$

Это и есть наибольшая скорость, с которой пользователь может надежно передавать сообщения при данных значениях Q, q, K и SNR, используя вышеописанный метод передачи. Следует подчеркнуть, что эта величина не тождественна пропускной способности канала «вверх», так как методы приема и передачи фиксированы.

Функционирование системы множественного доступа описываемого типа в условиях подавления с использованием сосредоточенной помехи

Рассмотрим систему множественного доступа описанного выше типа, использующую канал с аддитивным белым гауссовым шумом, в которой мощность передаваемых сигналов выбирается таким образом, чтобы мощность сигналов на приемном конце была одинакова и равна P.

Ниже будет рассмотрена ситуация, в которой враждебный пользователь использует для подавления сигналы, аналогичные полезным, т. е. имеющие такую же форму и занимающие полосу частот той же ширины, что и любой из полезных сигналов. Следует учесть, что формирование сигнала, антиподального подавляемому, является тех-

нически нереализуемым. (Так, в реальной системе, построенной в соответствии с вышеописанными принципами, в случае, когда X стремится подавить сигнал, который пользователь A передает пользователю B для того, чтобы сформировать соответствующий сигнал X, необходимо в частности удовлетворительно оценить параметры сразу трех различных каналов: AX, XB и AB.) Будем полагать, что сигнал, который формируется враждебным пользователем, имеет случайную фазу.

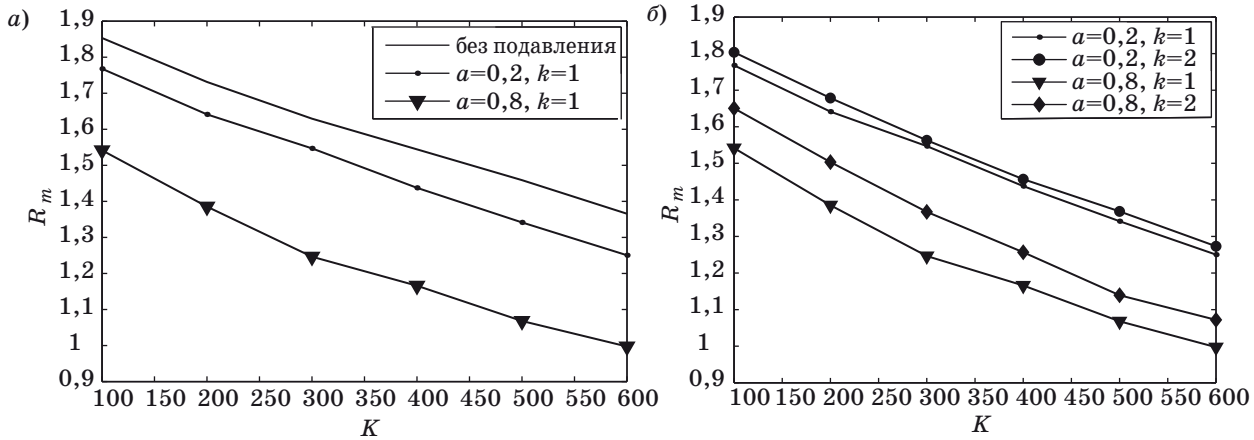
Заметим, что рассматриваемая стратегия удобна для враждебного пользователя, так как позволяет ему имитировать работу легальных пользователей и тем самым затрудняет обнаружение, и потому предлагаемая модель представляется вполне реалистичной. Вместе с тем в настоящей работе предполагается, что враждебный пользователь обладает энергетическими ресурсами большими, чем любой из легальных пользователей, т. е. как мощность каждого из сигналов, переданных враждебным пользователем, на приемном конце, так и количество сигналов, передаваемых рассматриваемым пользователем, могут быть больше (возможно, существенно больше), чем соответствующие значения для легального пользователя.

Будем считать, что враждебный пользователь способен определить подканалы, в которых ведется передача, за время τ ($\tau = gT$, $g << 1$) и сгенерировать $A = aK$ сигналов, передаваемая мощность которых такова, что мощность соответствующих сигналов на приемном конце оказывается равной $\tilde{P} = kP$, где P — мощность сигнала от легального пользователя на приемном конце. Прежде всего, рассмотрим случай (рисунок, *а*), когда $k = 1$, а количество сигналов в одном случае относительно невелико по сравнению с числом активных пользователей ($a = 0{,}2$), а во втором, напротив, сравнимо с числом активных пользователей ($a = 0{,}8$). Здесь и далее для моделирования будут использоваться следующие значения системных параметров: $Q = 4096$, $q = 4$, $SNR = 10$ дБ, $\tau = 0{,}1T$ и изменения значения числа мешающих пользователей в диапазоне значений от 100 до 600. Для сравнения приводятся значения R_m (в битах на OFDM-символ) для случая, когда подавление полезных сигналов не происходит.

Как видно из рисунка, увеличение числа подавляемых полезных сигналов A ожидаемо ведет к снижению значения максимальной скорости передачи. Тем не менее даже для сравнительно большого числа A уменьшение значения максимальной скорости передачи составляет не более 0,4 бита на один OFDM-символ (т. е. не более 20 % от максимально возможной для данного случая скорости 2 бита на один OFDM-символ). Для случая же, когда враждебный пользователь обладает существенно меньшими ресурсами (т. е. способен подавлять не более 20 % от общего числа активных пользователей), уменьшение скорости и вовсе составляет не более 0,1 бита на OFDM-символ (т. е. менее 5 % от максимально возможной скорости).

Рассмотрим теперь ситуацию (рисунок, *б*), в которой каждый из генерируемых враждебным пользователем сигналов имеет мощность большую, чем сигналы, которые используют легальные пользователи. Приведены графики для случаев $k = 1$ (мощность сигналов подавления равна мощности полезных сигналов) и $k = 2$ (мощность сигналов подавления в два раза больше мощности полезных сигналов).

Как видно из вышеприведенного графика, увеличение мощности не приводит к снижению максимально возможной скорости передачи данных. Напротив, при использовании такой техни-



■ *Зависимость максимально возможной скорости передачи от числа мешающих пользователей: а — при различном числе сигналов подавления и при отсутствии таковых; б — для различных значений числа подавляемых полезных сигналов и значений мощности сигналов преднамеренной помехи*

ки максимальная скорость передачи лишь увеличивается, причем когда сигналов подавления больше, рост этой величины оказывается более значительным, чем когда число сигналов, используемых для подавления, сравнительно невелико. Это, по-видимому, можно объяснить тем, что при сравнительно большом числе подканалов, в которые передаются сигналы подавления, вероятность того, что и сигнал от рассматриваемого пользователя окажется в числе подавляемых, возрастает. Следует отметить, что использование сигналов большей мощности лишь увеличивает среднюю мощность. Когда число сигналов сравнительно велико, это означает, что существенно снижается вероятность ошибки, так как для ошибки необходимо, чтобы порог в подканале, по которому передается полезный сигнал, не был превышен, а вероятность такого исхода уменьшается с ростом средней мощности передаваемых сигналов.

Заключение

Результаты, полученные в ходе имитационного моделирования системы описанного типа, позволяют сделать некоторые выводы относительно эффективности различных вариантов рассмотренной нами стратегии подавления. В целом, описанная выше система множественного доступа отличается хорошей устойчивостью к такой стратегии подавления, как намеренная передача сигналов, имитирующих сигналы легальных пользователей, в подканалах используемых легальными пользователями для передачи. Сравнение различных вариантов реализации этой стратегии с точки зрения снижения эффективности передачи данных легальными пользователями свидетельствует о том, что увеличение мощности передаваемых сигналов нерационально, поскольку не приводит к более эффективному подавлению. Анализ результатов моделирования показывает, что более рациональным способом использования энергетического преимущества для целей подавления является перераспределение энергии для увеличения числа подканалов, в которых происходит подавление, однако даже использование значительного числа сигналов подавления не приводит к существенному ухудшению эффективности передачи данных.

Литература

1. **Зяблов В., Осипов Д.** Об оптимальном выборе порога в системе множественного доступа, основанной на перестроении ортогональных частот // Проблемы передачи информации / ИППИ РАН. 2008. Т. 44. Вып. 2. С. 23–31.
2. **Osipov D.** On the probabilistic description of a FH-OFDMA with a MAXP receiver // Problems of redundancy in information and control systems / Proc. of the XII Symp., St.-Petersburg, May 26–30, 2009. P. 144–149.
3. **Грошев Ф. В., Осипов Д. С.** Исследование пропускной способности системы множественного доступа с пороговым приемом // Информационные технологии и системы: Сб. 32-й конф. молодых ученых и специалистов ИППИ РАН, Бекасово, Россия, 15–18 декабря 2009 г. С. 152–155.
4. **Галлагер Р.** Теория информации и надежная связь. — М.: Сов. радио, 1974. — 720 с.

УДК 004.728.3.057.4

ОЦЕНКА ПРОИЗВОДИТЕЛЬНОСТИ ПРОСТЕЙШЕЙ СИСТЕМЫ АБОНЕНТСКОЙ КООПЕРАЦИИ

С. Д. Андреев,
канд. техн. наук, научный сотрудник
А. В. Винель,
канд. техн. наук, старший научный сотрудник
О. С. Галинина,
соискатель
Санкт-Петербургский институт информатики и автоматизации РАН

Рассмотрена модель системы абонентской кооперации, включающая трех абонентов. Получены замкнутые выражения для средней задержки передачи сообщения, а также для пропускной способности, энергетического потребления и энергетической эффективности абонентов-источников. Точность найденных характеристик подтверждается при помощи имитационного моделирования.

Ключевые слова *— сотовая сеть, абонентская кооперация, система массового обслуживания, энергетическая эффективность.*

Введение

Беспроводные сети передачи информации получают широкое распространение, обусловленное скорым выходом в свет новейших телекоммуникационных протоколов [1, 2]. Будущее развитие систем беспроводной связи во многом зависит от того, насколько успешно будет преодолен дисбаланс между требуемым качеством обслуживания и ограниченным спектральным ресурсом системы связи. Вместе с тем задача повышения *спектральной* эффективности системы понемногу уступает место задаче повышения ее *энергетической* эффективности в первую очередь для малогабаритных мобильных устройств в силу увеличивающегося разрыва между доступной и требуемой емкостью их аккумуляторной батареи [3].

Эффективное управление ресурсом системы передачи информации становится особенно важным для технологий, в которых множество абонентов делят между собой ограниченный спектральный ресурс [4]. В настоящее время «уровневый» подход к построению системы передачи информации доминирует при разработке сетевых решений, причем каждый уровень рассматривается независимо, поддерживая таким образом механизм абстрагирования. Среди таких уровней *физический* отвечает за организацию битовой передачи, тогда как уровень *управления доступом к среде* (УДС) регламентирует доступ абонентов к общему ресурсу системы связи.

Однако традиционная «уровневая» архитектура оказывается недостаточно гибкой и приводит к неэффективному использованию ресурса [5]. Для преодоления данного ограничения требуется новый интегральный и адаптивный подход. Как следствие, кросс-уровневые подходы, совместно учитывающие физический уровень и уровень УДС, получают все больше внимания исследователей [4]. Для достижения кросс-уровневых преимуществ разрабатываются решения со «знанием о канале», которые в явном виде используют информацию о состоянии беспроводной среды передачи. Они, как правило, реализуют дополнительные возможности взаимодействия физического уровня и уровня УДС, повышая тем самым адаптивность к меняющимся требованиям по качеству обслуживания, входному потоку, а также свойствам канала связи [6–8].

По мере того как возрастает степень мобильности абонентов, фокус исследовательских работ в данной области смещается с пропускной способности [9] в сторону учета энергетического потребления на всех уровнях беспроводной сети связи [10] — от архитектуры [11] до применяемых алгоритмов [12]. В последнее время особый интерес вызывают *кооперативные* кросс-уровневые подходы [13, 14], позволяющие получить выигрыш

за счет использования различия в характеристиках абонентских каналов связи и, как следствие, повысить производительность системы передачи информации.

В то время как все большее число абонентов делят между собой единый спектральный ресурс, а современные *сотовые* системы связи переходят к более «агрессивным» способам использования частотного диапазона [1, 2], интерференция становится одним из важнейших факторов, сдерживающих рост производительности систем передачи информации. Поскольку передача в беспроводном канале связи принципиально является широковещательной, сообщение от одного абонента накладывается на сообщения от других абонентов, снижая энергетическую эффективность системы в целом. Однако абоненты могут повысить свою энергетическую эффективность, функционируя кооперативно [15, 16]. Такое пространственное управление ресурсом приобретает все большую важность для улучшения работы абонентов, находящихся на удалении от центра соты [5].

С другой стороны, кооперативное функционирование приводит к повышенному потреблению энергии за счет передачи дополнительных сообщений и, возможно, к росту задержки передачи информации из-за промежуточных пересылок. Однако увеличение задержки может быть использовано при управлении скоростью передачи абонента, поскольку понижение скорости зачастую влечет повышение энергетической эффективности. В итоге, представляется необходимым исследовать все базовые обменные соотношения, связанные с кооперативным функционированием, и выявить условия, в которых оно приводит к реальному росту производительности беспроводной системы передачи информации.

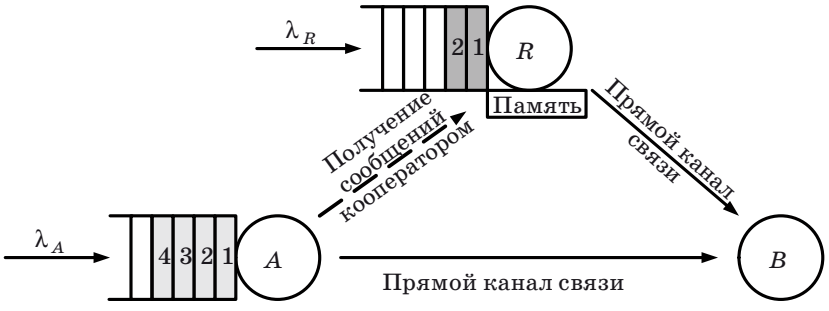
В настоящее время значительный интерес для повышения производительности сетей передачи информации представляет кооперация между соседними абонентами системы связи. Поскольку энергетические затраты на обеспечение надежной передачи возрастают экспоненциально с увеличением расстояния [17], предпочтительно отправлять данные, используя несколько промежуточных звеньев [18]. По этой причине *абонентская кооперация* может стать одной из ключевых технологий для повышения энергетической эффективности современных сотовых систем передачи информации.

При использовании кооперации необходимо исключить сценарии, приводящие к недостаточному росту производительности абонента-отправителя сообщений [19]. В результате задача эффективного выбора кооператоров сводится к учету обменного соотношения между повышением характеристик отправителя и дополнительными затратами кооператора. В данной работе приводится анализ производительности простейшей, но в то же время реалистичной системы абонентской кооперации, которая далее именуется *базовой*. Получена оценка средней задержки передачи сообщения для всех абонентов-источников в рамках данной модели. Кроме того, выполнен расчет пропускной способности, энергетического потребления и энергетической эффективности абонентов-источников.

Модель системы и обозначения

1. *Основные допущения*.

Рассмотрим систему передачи информации с возможностью абонентской кооперации, имеющую простейшую топологию, которая включает два абонента-источника (рис. 1) и одного получателя сообщений. Будем далее называть абонента *A отправителем*. У отправителя возникают новые сообщения со средней интенсивностью λ_A. Абонент R далее называется *кооператором*. У кооператора возникают новые сообщения со средней интенсивностью λ_R. Кроме того, кооператор имеет возможность принимать сообщения отправителя и сохранять их для последующей передачи. Абонент B, называемый *базовой станцией* (БС), принимает сообщения отправителя и кооператора. Ниже приведены основные допущения, составляющие модель системы.



■ ***Рис. 1.*** *Базовая модель кооперативной системы*

Допущение 1. Время работы системы передачи информации дискретно. Единица системного времени называется *слотом*. Все передаваемые сообщения имеют одинаковую длину. Передача одного сообщения от источника получателю занимает ровно один слот.

Допущение 2. Количества новых сообщений, поступающих к отправителю и к кооператору в течение одного слота, являются независимыми и одинаково распределенными (i.i.d) случайными величинами со средними значениями λ_A и λ_R соответственно. Для простоты последующего анализа в данной работе предполагается *пуассоновский* входной поток сообщений. БС не имеет собственного исходящего трафика.

Допущение 3. Отправитель и кооператор имеют очереди неограниченной длины для хранения собственных сообщений. Кроме того, у кооператора имеется дополнительная ячейка памяти для хранения *единственного* сообщения, полученного от отправителя в целях последующей кооперативной передачи. Достаточность одной ячейки памяти для хранения принятых от отправителя сообщений будет показана ниже.

Допущение 4. Система централизована и управляется БС, где имеется справедливый *стохастический* кольцевой планировщик, который чередует абонентов, осуществляющих доступ к каналу связи, с равной вероятностью (см. рис. 2 в качестве примера функционирования системы, приведенной на рис. 1, без дополнительных поступлений). В частности, если как у отправителя, так и у кооператора имеются готовые для передачи сообщения, один из них получает следующий слот для передачи с вероятностью 0,5, а другой простаивает. Если один из абонентов-источников не имеет готовых для передачи сообщений, второй получает следующий слот для передачи с вероятностью 1. В том случае, если оба источника не имеют готовых для передачи сообщений, система простаивает. Информация о расписании передач поступает по выделенному каналу связи и не потребляет ресурс системы.

Допущение 5. Канал связи подвержен ошибкам и описывается моделью канала связи с многопакетным приемом [20]. Передаваемое сообщение искажается с некоторой постоянной вероятностью, которая зависит только от типа канала связи и количества одновременно передающих абонентов.

Параметры модели следующие:

$$p_{AB} = \Pr\left\{\begin{matrix}\textit{сообщение от } A \textit{ принято на } B \\ \textit{передает только } A\end{matrix}\right\};$$

$$p_{RB} = \Pr\left\{\begin{matrix}\textit{сообщение от } R \textit{ принято на } B \\ \textit{передает только } R\end{matrix}\right\};$$

$$p_{AR} = \Pr\left\{\begin{matrix}\textit{сообщение от } A \textit{ принято на } R \\ \textit{передает только } A\end{matrix}\right\};$$

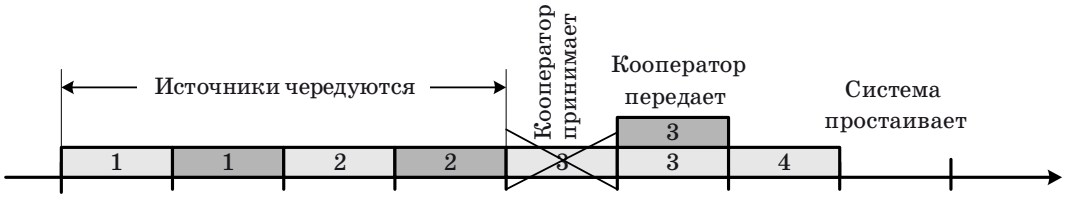
$$p_{CB} = \Pr\left\{\begin{matrix}\textit{сообщение от } A \textit{ принято на } B \\ A \textit{ и } R \textit{ кооперируют}\end{matrix}\right\}.$$

Предполагается, что $p_{AR} > p_{AB}$, а $p_{CB} > p_{AB}$.

Информация обратной связи относительно успеха передач абонентов поступает по выделенному каналу связи и не потребляет ресурс системы. Она показывает, было ли сообщение успешно принято на БС к концу текущего слота. В случае если сообщение искажено, оно передается источником повторно. Максимальное количество возможных попыток передачи сообщения не ограничено. Кооператор не может передавать и принимать сообщения одновременно.

Допущение 6. При *первой* попытке передачи сообщения отправителем кооператор осуществляет попытку его приема с вероятностью 1. Согласно допущению 5, сообщение успешно принимается кооператором с вероятностью p_{AR}. В случае успеха кооператор сохраняет принятое сообщение в дополнительной ячейке памяти, уничтожая предыдущее ее содержимое. Также согласно допущению 5, сообщение отправителя успешно принимается БС с вероятностью p_{AB}. В случае неуспеха отправитель передает сообщение повторно.

Допущение 7. При *повторной* попытке передачи сообщения отправителем кооператор осуществляет одно из следующих действий. Если передаваемое сообщение уже сохранено в дополнительной ячейке памяти, кооператор передает его *одновременно* с отправителем с вероятностью 1 (см. рис. 2). Согласно допущению 5, сообщение успешно принимается БС с вероятностью p_{CB}. В противном случае кооператор снова осуще-



■ ***Рис. 2.*** *Пример работы системы абонентской кооперации*

ствляет попытку приема сообщения отправителя с вероятностью 1 (см. допущение 6).

Отметим, что согласно допущениям 6 и 7, одной дополнительной ячейки памяти кооператора для хранения принятых от отправителя сообщений оказывается достаточно для работы системы абонентской кооперации. При этом отправитель не имеет информации относительно кооперативной помощи, и кооператор не высылает ему подтверждений, в отличие от подхода в работе [20]. Кооператор повышает пропускную способность отправителя за счет снижения собственной энергетической эффективности. Дополнительная энергия расходуется кооператором при получении сообщений отправителя и при их передаче одновременно с отправителем.

Кооператор может в ряде случаев не принимать сообщения отправителя или не передавать их, следуя некоторой политике кооперативного функционирования. В данной работе такие случаи *избирательной* кооперации не рассматриваются, и изложение сужается до сценария, когда кооператор осуществляет обязательную попытку приема всех сообщений отправителя и передает их одновременно с отправителем каждый раз, когда они имеются в дополнительной ячейке памяти (см. допущения 6 и 7).

2. *Основные обозначения.*

Предлагаемый аналитический подход направлен на изучение времени обслуживания сообщения. *Временем обслуживания* назовем интервал времени от момента, когда сообщение готово к обслуживанию, до момента окончания его обслуживания [21].

Обозначим время обслуживания сообщения от A через $T_{AR}(\lambda_A,\lambda_R) \triangleq T_{AR}$, где знак ‘$\triangleq$’ используется в смысле «равно по обозначению». Кроме того, введем в рассмотрение *среднее* время обслуживания сообщения от A как $\tau_{AR}(\lambda_A,\lambda_R) \triangleq \tau_{AR} = E[T_{AR}]$. Далее обозначим через $\tau_{AR}(\lambda_A,0) \triangleq \tau_{A0}$ среднее время обслуживания сообщения от A при условии, что $\lambda_R = 0$.

Симметрично введем в рассмотрение время обслуживания сообщения от R как $T_{RA}(\lambda_R,\lambda_A) \triangleq T_{RA}$ и соответствующую среднюю длительность как $\tau_{RA}(\lambda_R,\lambda_A) \triangleq \tau_{RA} = E[T_{RA}]$. Аналогично условное среднее время обслуживания составит $\tau_{RA}(\lambda_R,0) \triangleq \tau_{R0}$ при $\lambda_A = 0$.

Отметим, что как при наличии, так и при отсутствии кооперации в силу геометрического распределения величины T_{R0} выполняется следующее соотношение:

$$\tau_{R0} = \frac{1}{p_{RB}}, \qquad (1)$$

тогда как только для системы без кооперации выполняется

$$\tau_{A0} = \frac{1}{p_{AB}}. \qquad (2)$$

Нахождение τ_{A0} в случае кооперативного функционирования является более сложной задачей и будет произведено ниже.

Обозначим количества сообщений в очередях абонентов A и R к началу некоторого слота t через $Q_A^{(t)}$ и $Q_R^{(t)}$ соответственно. Поскольку далее будет исследоваться система абонентской кооперации в стационарном состоянии, верхний индекс t у величин $Q_A^{(t)}$ и $Q_R^{(t)}$ будем опускать.

Обозначим *коэффициент загрузки* [22] абонента A через $\rho_{AR}(\lambda_A,\lambda_R) \triangleq \rho_{AR}$. По определению:

$$\rho_{AR} = \Pr\{Q_A \neq 0\} = \lambda_A \tau_{AR}. \qquad (3)$$

В частности, коэффициент загрузки абонента A при условии $\lambda_R = 0$ может быть установлен как $\rho_{AR}(\lambda_A,0) \triangleq \rho_{A0} = \lambda_A \tau_{A0}$. С учетом (2) для системы без кооперации ρ_{A0} дополнительно упрощается до $\rho_{A0} = \dfrac{\lambda_A}{p_{AB}}$.

По аналогии коэффициент загрузки абонента R обозначим через $\rho_{RA}(\lambda_R,\lambda_A) \triangleq \rho_{RA}$. Также по определению:

$$\rho_{RA} = \Pr\{Q_R \neq 0\} = \lambda_R \tau_{RA}. \qquad (4)$$

Симметрично коэффициент загрузки абонента R при условии $\lambda_A = 0$ может быть найден как $\rho_{RA}(\lambda_R,0) \triangleq \rho_{R0} = \lambda_R \tau_{R0}$. С учетом (1) для обеих систем с кооперацией и без кооперации ρ_{R0} упрощается до $\rho_{R0} = \dfrac{\lambda_R}{p_{RB}}$. Основные обозначения сведены в таблицу.

Оценка характеристик системы

1. *Общие утверждения.*

Рассмотрим очередь абонента A. Напомним, что, по определению, $\rho_{AR} = \Pr\{Q_A \neq 0\}$ и положим $\rho_{A0} > \rho_{R0}$ для определенности. Можно сформулировать следующее утверждение.

Утверждение 1. Для коэффициента загрузки абонента A при любых λ_A и λ_R справедливо соотношение

$$\rho_{AR} \leq \frac{\rho_{A0}}{1-\rho_{R0}}. \qquad (5)$$

Другое важное утверждение можно сформулировать, исходя из рассмотрения условия нормировки порождающей функции системы или балансных уравнений соответствующей вложенной цепи Маркова.

Утверждение 2. Для коэффициентов загрузки абонентов A и R при любых λ_A и λ_R справедливо соотношение

■ *Принятые обозначения*

Обозначение	Описание параметра
λ_A	Средняя интенсивность входного потока сообщений к абоненту A
λ_R	Средняя интенсивность входного потока сообщений к абоненту R
p_{AB}	Вероятность успешного приема от A на B при передаче от A
p_{RB}	Вероятность успешного приема от R на B при передаче от R
p_{AR}	Вероятность успешного приема от A на R при передаче от A
p_{CB}	Вероятность успешного приема от A на B при передаче A и R
τ_{AR}	Среднее время обслуживания сообщения от A
τ_{RA}	Среднее время обслуживания сообщения от R
ρ_{AR}	Коэффициент загрузки абонента A
ρ_{RA}	Коэффициент загрузки абонента R
q_A	Средняя длина очереди абонента A
q_R	Средняя длина очереди абонента R
δ_A	Средняя задержка передачи сообщения от A
δ_R	Средняя задержка передачи сообщения от R
η_A	Средняя интенсивность выходного потока сообщений от A
η_R	Средняя интенсивность выходного потока сообщений от R
ε_A	Среднее энергетическое потребление абонента A
ε_R	Среднее энергетическое потребление абонента R
φ_A	Средняя энергетическая эффективность абонента A
φ_R	Средняя энергетическая эффективность абонента R

$$\rho_{AR} - \rho_{RA} = \rho_{A0} - \rho_{R0}. \qquad (6)$$

Доказательства утверждений 1 и 2 не включены в текст данной работы в силу своего значительного объема.

Утверждение 3. Для коэффициента загрузки абонента R при любых λ_A и λ_R справедливо соотношение

$$\rho_{RA} = \rho_{AR} - \rho_{A0} + \rho_{R0} \le \frac{\rho_{A0}}{1-\rho_{R0}} - \rho_{A0} + \rho_{R0}. \qquad (7)$$

Доказательство утверждения 3 немедленно следует из (5) и (6).

Полученные оценки (5) и (7) справедливы для системы как без кооперации, так и с кооперацией. Далее рассматривается система без кооперации, а затем предложенный подход к оценке средней задержки обобщается на случай системы с кооперацией.

2. *Система без кооперации.*

Опишем работу абонента A в терминах теории массового обслуживания. Для этого рассмотрим систему массового обслуживания, связанную с абонентом A. В силу зависимости между очередями абонентов A и R известная формула Полячека–Хинчина [22] не дает точного значения для числа сообщений в очереди абонента A. Будем, тем не менее, использовать эту формулу для получения приближенного значения средней длины очереди абонента A как

$$q_A \cong \lambda_A E[T_{AR}] + \frac{\lambda_A^2 E[T_{AR}^2]}{2(1-\lambda_A E[T_{AR}])} = \lambda_A \tau_{AR} + \frac{\lambda_A^2 E[T_{AR}^2]}{2(1-\lambda_A \tau_{AR})}, \qquad (8)$$

где $\tau_{AR} = E[T_{AR}]$ представляет собой среднее время обслуживания сообщений от A (первый момент для времени обслуживания T_{AR}), а $E[T_{AR}^2]$ представляет собой второй момент для времени обслуживания. Учитывая (3) и (8), имеем

$$q_A \cong \rho_{AR} + \frac{\lambda_A^2 E[T_{AR}^2]}{2(1-\rho_{AR})}. \qquad (9)$$

Покажем теперь, как вычисляются отдельные компоненты выражения (9). Рассмотрим время обслуживания некоторого сообщения от A, которое начинается в момент времени, когда сообщение готово к обслуживанию, и заканчивается в момент окончания обслуживания. Напомним, что планировщик на БС является стохастическим, т. е. назначает очередной слот абоненту A с вероятностью 0,5, если оба абонента-источника загружены. Таким образом, в каждом слоте, для которого $Q_R \neq 0$ и $Q_A \neq 0$, сообщение от A включается в расписание с вероятностью 0,5. Введем следующую вспомогательную вероятность:

$$\gamma_A \triangleq \Pr\{Q_R \neq 0 \,|\, Q_A \neq 0\} = \frac{\Pr\{Q_R \neq 0, Q_A \neq 0\}}{\Pr\{Q_A \neq 0\}}.$$

Очевидно, что планировщик назначит очередной слот абоненту R с вероятностью $0{,}5\gamma_A$ и назначит слот абоненту A с дополнительной вероятностью $1 - 0{,}5\gamma_A$.

Рассмотрим вероятность события, когда $Q_R \neq 0$ и $Q_A \neq 0$ одновременно. По формуле полной веро-

ятности: $\Pr\{Q_R \neq 0,\ Q_A \neq 0\} = \Pr\{Q_R \neq 0\} - \Pr\{Q_R \neq 0,\ Q_A = 0\}$. С другой стороны, по определению: $\Pr\{Q_R \neq 0\} = \rho_{RA}$. Далее, для вероятности $\Pr\{Q_R \neq 0, Q_A = 0\}$ можно аналогично записать выражение $\Pr\{Q_R \neq 0, Q_A = 0\} = \Pr\{Q_A = 0\} - \Pr\{Q_R = 0,\ Q_A = 0\}$. Используя определение ρ_{AR}, отметим, что $\Pr\{Q_A = 0\} = 1 - \rho_{AR}$. Кроме того, имеем $\Pr\{Q_R = 0, Q_A = 0\} = 1 - \rho_{A0} - \rho_{R0}$.

Суммируя предыдущие рассуждения, имеем $\Pr\{Q_R \neq 0,\ Q_A \neq 0\} = \rho_{AR} + \rho_{RA} - \rho_{A0} - \rho_{R0}$. Кроме того, из утверждения 2 следует, что $\Pr\{Q_R \neq 0,\ Q_A \neq 0\} = 2(\rho_{AR} - \rho_{A0})$. В итоге получаем, что

$$0{,}5\gamma_A = 0{,}5 \cdot \frac{\Pr\{Q_R \neq 0,\ Q_A \neq 0\}}{\Pr\{Q_A \neq 0\}} = 1 - \frac{\rho_{A0}}{\rho_{AR}}.$$

Можно выписать следующее распределение для времени обслуживания сообщений от A:

$$\Pr\{T_{AR} = n\} = p_{AB}(1 - 0{,}5\gamma_A) \times \\ \times (1 - p_{AB}(1 - 0{,}5\gamma_A))^{n-1}.$$

Данное выражение учитывает, что из n слотов, потраченных на передачу сообщения от A, последний слот был назначен абоненту A, и передача в нем была успешной. Предыдущие n–1 слотов либо не были назначены абоненту A, либо передача в них не была успешной.

Вычисляя первый и второй моменты для времени обслуживания $E[T_{AR}]$ и $E[T_{AR}^2]$, а также используя (9) и формулу Литтла в виде $q_A = \lambda_A \delta_A$, легко выразить среднюю задержку передачи сообщения абонентом A как

$$\delta_A \cong \frac{\rho_{AR}}{\lambda_A} + \frac{\lambda_A(2 - p_{AB}(1 - 0{,}5\gamma_A))}{2(1 - \rho_{AR})p_{AB}^2(1 - 0{,}5\gamma_A)^2}.$$

Характеристики абонента R вычисляются аналогично в силу симметричности каналов связи. С учетом

$$0{,}5\gamma_R = 1 - \frac{\rho_{R0}}{\rho_{RA}},$$

где $\gamma_R \triangleq \dfrac{\Pr\{Q_R \neq 0,\ Q_A \neq 0\}}{\Pr\{Q_R \neq 0\}}$, средняя задержка передачи сообщения абонентом R составляет

$$\delta_R \cong \frac{\rho_{RA}}{\lambda_R} + \frac{\lambda_R(2 - p_{RB}(1 - 0{,}5\gamma_R))}{2(1 - \rho_{RA})p_{RB}^2(1 - 0{,}5\gamma_R)^2}.$$

Предлагаемый подход к анализу системы абонентской кооперации позволяет также установить точные значения для средней интенсивности выходного потока сообщений (пропускной способности) от абонента A и R. В частности, пропускная способность абонента A составляет

$$\eta_A = \begin{cases} \lambda_A, \text{ нет насыщения} \\ \dfrac{1 - \lambda_R\tau_{R0}}{\tau_{A0}}, \text{ насыщение для } A. \\ \dfrac{1}{2\tau_{A0}}, \text{ насыщение для } A, R \end{cases}$$

Аналогично пропускная способность абонента R вычисляется как

$$\eta_R = \begin{cases} \lambda_R, \text{ нет насыщения} \\ \dfrac{1 - \lambda_A\tau_{A0}}{\tau_{R0}}, \text{ насыщение для } R. \\ \dfrac{1}{2\tau_{R0}}, \text{ насыщение для } A, R \end{cases}$$

Данные выражения можно дополнительно упростить с учетом соотношений (1) и (2). Здесь условия насыщения определяются следующим образом:

- насыщение для A: $(\lambda_A\tau_{A0} + \lambda_R\tau_{R0} > 1)$, и при этом $(\lambda_R\tau_{R0} < 0{,}5)$;
- насыщение для R: $(\lambda_A\tau_{A0} + \lambda_R\tau_{R0} > 1)$, и при этом $(\lambda_A\tau_{A0} < 0{,}5)$;
- насыщение для A и R: $(\lambda_A\tau_{A0} > 0{,}5)$, и при этом $(\lambda_R\tau_{R0} > 0{,}5)$.

Кроме того, можно установить среднее энергетическое потребление абонента A как

$$\varepsilon_A = P_{TX}\,\eta_A\tau_{A0} + P_I\,(1 - \eta_A\tau_{A0})$$

и среднее энергетическое потребление абонента R как

$$\varepsilon_R = P_{TX}\,\eta_R\tau_{R0} + P_I\,(1 - \eta_R\tau_{R0}).$$

Здесь P_{TX} — средняя мощность, затрачиваемая абонентом в состоянии передачи сообщений, а P_I — средняя мощность, затрачиваемая абонентом в состоянии простоя. Тогда средняя энергетическая эффективность абонентов A и R немедленно устанавливается с помощью выражений $\varphi_A = \dfrac{\eta_A}{\varepsilon_A}$ и $\varphi_R = \dfrac{\eta_R}{\varepsilon_R}$ соответственно.

3. *Система с кооперацией.*

Для описания системы с кооперацией вначале рассмотрим важный частный случай, когда очередь абонента R всегда пуста. Выразим распределение числа слотов, необходимых для обслуживания некоторого сообщения от абонента A. Затем используем найденное распределение, чтобы обобщить предлагаемый подход на случай непустой очереди R. Все соответствующие характери-

стики для системы с кооперацией пометим символом '*' в верхнем регистре.

Случай I. Очередь абонента R всегда пуста ($\lambda_R = 0$).

Проводя рассуждения, аналогичные рассуждениям из предыдущего подраздела, выразим искомое время обслуживания сообщения от A как

$$\Pr\{T^*_{A0} = n\} = X(1 - p_{CB})^{n-1} - Y[(1 - p_{AB})(1 - p_{AR})]^{n-1},$$

где $X = \dfrac{p_{AR}(1 - p_{AB})p_{CB}}{1 - p_{CB} - (1 - p_{AB})(1 - p_{AR})}$ и $Y = X - p_{AB}$.

Переходя к среднему времени обслуживания, имеем

$$\tau^*_{A0} = \frac{p_{CB} + (1 - p_{AB})p_{AR}}{p_{CB}[p_{AB} + (1 - p_{AB})p_{AR}]}. \qquad (10)$$

Случай II. Очередь абонента R не всегда пуста ($\lambda_R > 0$).

Обобщим приведенный выше подход на наиболее сложный кооперативный случай с $\lambda_R > 0$. Опуская трудоемкие, но очевидные преобразования, для времени обслуживания сообщения от A получаем

$$\Pr\{T^*_{AR} = n\} = X(1 - 0{,}5\gamma^*_A) \times (1 - p_{CB}(1 - 0{,}5\gamma^*_A))^{n-1} - Y(1 - 0{,}5\gamma^*_A) \times (1 - p_A(1 - 0{,}5\gamma^*_A))^{n-1},$$

где $0{,}5\gamma^*_A = 1 - \dfrac{\rho^*_{A0}}{\rho^*_{AR}}$, а также для краткости $p_A = p_{AB} + p_{AR} - p_{AB} \cdot p_{AR}$. Здесь p_A — вероятность успешной доставки сообщения от A к R или на БС.

Коэффициенты загрузки абонентов A и R (ρ^*_{AR} и ρ^*_{RA}) могут быть вычислены по аналогии с соответствующими величинами в системе без кооперации с учетом того факта, что $\rho^*_{A0} \triangleq \lambda_A \tau^*_{A0}$, где выражение для τ^*_{A0} задается формулой (10).

Наконец, вычисляя второй момент для найденного времени обслуживания, получаем искомое выражение для средней задержки передачи сообщения абонентом A как

$$\delta^*_A \cong \frac{\rho^*_{AR}}{\lambda_A} + \frac{\lambda_A}{2(1 - \rho^*_{AR})(1 - 0{,}5\gamma^*_A)^2} \times \left[X \cdot \frac{2 - p_{CB}(1 - 0{,}5\gamma^*_A)}{p^3_{CB}} - Y \cdot \frac{2 - p_A(1 - 0{,}5\gamma^*_A)}{p^3_A} \right],$$

где X и Y определены выше.

Итоговые значения средней задержки передачи сообщения δ^*_R абонента R, а также пропускной способности η^*_A и η^*_R абонентов A и R в системе с кооперацией аналогичны соответствующим характеристикам этого абонента в системе без кооперации, вычисленным ранее, с учетом выражения (10). Аналогично среднее энергетическое потребление абонента A в рассматриваемом случае составит

$$\varepsilon^*_A = P_{TX}\,\eta^*_A \tau^*_{A0} + P_I\,(1 - \eta^*_A \tau^*_{A0}),$$

тогда как среднее энергетическое потребление абонента R

$$\varepsilon^*_R = P_{TX}\left(\eta^*_R \tau_{R0} + \eta^*_A \cdot \frac{1 - p_{AB}\tau^*_{A0}}{p_{CB} - p_{AB}}\right) + P_{RX}\left(1 - \eta^*_R \tau_{R0} - \eta^*_A \cdot \frac{1 - p_{AB}\tau^*_{A0}}{p_{CB} - p_{AB}}\right),$$

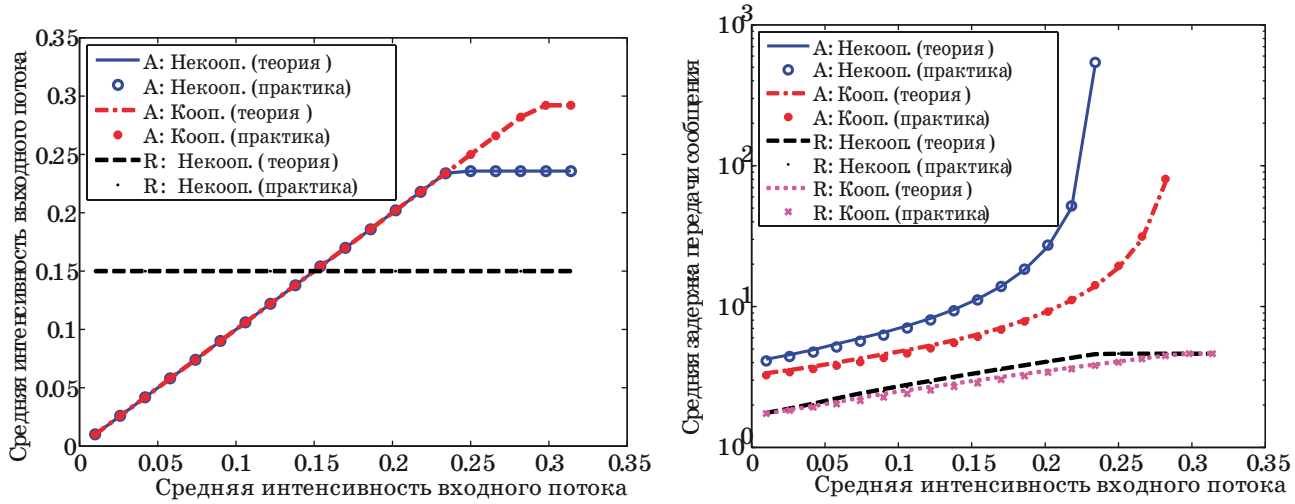
где P_{RX} — средняя мощность, затрачиваемая абонентом в состоянии приема сообщений. Средняя энергетическая эффективность абонентов A и R по-прежнему устанавливается с помощью выражений $\varphi^*_A = \dfrac{\eta^*_A}{\varepsilon^*_A}$ и $\varphi^*_R = \dfrac{\eta^*_R}{\varepsilon^*_R}$ соответственно.

Заключение

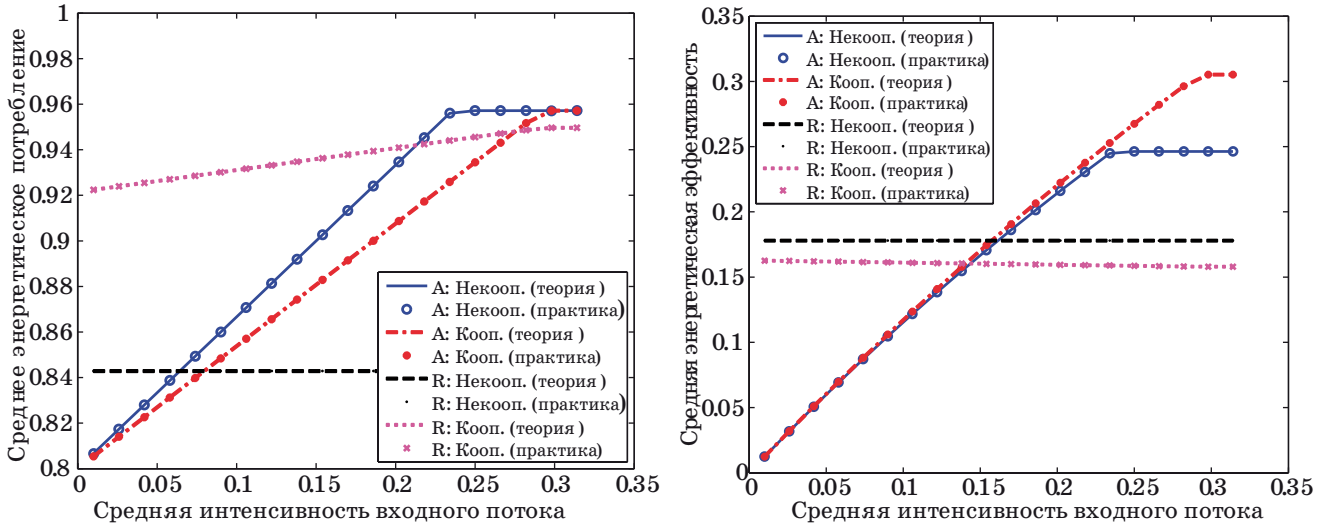
Результаты имитационного моделирования системы передачи информации с возможностью абонентской кооперации, в частности значения пропускной способности, задержки передачи сообщения, энергетического потребления и энергетической эффективности абонентов, представлены на рис. 3 и 4. Параметры имитационного моделирования установлены следующие: $p_{AB} = 0{,}3$, $p_{RB} = 0{,}7$, $p_{AR} = 0{,}4$, $p_{CB} = 0{,}5$, $\lambda_R = 0{,}15$ [20], а λ_A варьируется в диапазоне, обеспечивающем стабильную работу системы.

Видно, что предлагаемый аналитический подход к оценке производительности простейшей системы абонентской кооперации позволяет получить характеристики изучаемой системы передачи информации с высокой точностью. Полученные результаты обосновывают выводы о количественных выигрышах при использовании абонентской кооперации. В частности, увеличение пропускной способности отправителя достигает 24 % в насыщении. Это приводит к заключению о перспективности технологии абонентской кооперации в современных беспроводных системах передачи информации.

В отличие от имеющихся в литературе подходов, где предпринимаются попытки лишь частично решить поставленные задачи, в данной работе применяется не только имитационное моделирование, но также разрабатываются формальные математические модели. Технология абонентской кооперации, исследованная авторами, может быть внедрена в аппаратные средства связи, производимые компаниями Motorola,



■ ***Рис. 3.*** *Зависимость пропускной способности (слева) и задержки передачи сообщения (справа) от интенсивности входного потока сообщений* λ_A



■ ***Рис. 4.*** *Зависимость энергетического потребления (слева) и энергетической эффективности (справа) абонентов от интенсивности входного потока сообщений* λ_A

Intel, Nokia и др. Рассмотренные в работе алгоритмы могут быть сертифицированы на совместимость с разрабатываемым протоколом беспроводных сетей доступа нового поколения IEEE 802.16m [1].

Работа выполнена при финансовой поддержке Правительства Санкт-Петербурга, а также при поддержке РФФИ по проектам № 10-08-01071-а и № 08-08-00403-а и в рамках программы фундаментальных исследований ОНИТ РАН по проекту 2.3.

Литература

1. **IEEE** Std 802.16m (D9). Amendment to IEEE Standard for Local and Metropolitan Area Networks. Part 16. Air Interface for Broadband Wireless Access Systems — Advanced Air Interface. http://ieee802.org/16/pubs/80216m.html (дата обращения: 11.11.2010).
2. **LTE** Release 10 & beyond (LTE-Advanced).
3. **Lahiri K., Raghunathan A., Dey S., Panigrahi D.** Battery-driven system design: A new frontier in low power design// Proc. Intl. Conf. on VLSI Design. Bangalore, India. Jan. 2002. P. 261–267.
4. **Andreev S.** et al. Active-Mode Power Optimization in OFDMA-Based Wireless Networks//Proc. IEEE BWA Workshop of Globecom. 2010. — 1 electron. opt. disk (CD-ROM).

5. **Andreev S., Galinina O., Vinel A.** Cross-Layer Channel-Aware Approaches for Modern Wireless Networks // Lecture Notes in Computer Science. Springer, 2010. Vol. 6235/2010. P. 163–179.
6. **Song G.** Cross-Layer Optimization for Spectral and Energy Efficiency: PhD thesis/School of Electrical and Computer Engineering; Georgia Institute of Technology, 2005. — 141 p.
7. **Miao G.** Cross-Layer Optimization for Spectral and Energy Efficiency: PhD thesis / School of Electrical and Computer Engineering; Georgia Institute of Technology, 2008. — 160 p.
8. **Kim H.** Exploring Tradeoffs in Wireless Networks under Flow-Level Traffic: Energy, Capacity and QoS: PhD thesis/University of Texas at Austin, 2009. — 162 p.
9. **Song G., Li Y.** Asymptotic throughput analysis for channel-aware scheduling // IEEE Trans. Commun. Oct. 2006. Vol. 54. N 10. P. 1827–1834.
10. **Anisimov A., Andreev S., Galinina O., Turlikov A.** Comparative Analysis of Sleep Mode Control Algorithms for Contemporary Metropolitan Area Wireless Networks // Lecture Notes in Computer Science. Springer, 2010. Vol. 6294/2010. P. 184–195.
11. **Benini L., Bogliolo A., de Micheli G.** A survey of design techniques for system-level dynamic power management // IEEE Trans. VLSI Syst. Jun. 2000. Vol. 8. P. 299–316.
12. **Schurgers C.** Energy-Aware Wireless Communications: PhD thesis/University of California Los Angeles, 2002. — 137 p.
13. **Pyattaev A., Andreev S., Vinel A., Sokolov B.** Client relay simulation model for centralized wireless networks // Proc. EUROSIM Congress, 2010. — 1 electron. opt. disk (CD-ROM).
14. **Pyattaev A., Andreev S., Koucheryavy Y., Moltchanov D.** Some Modeling Approaches for Client Relay Networks // Proc. IEEE CAMAD Workshop. — 1 electron. opt. disk (CD-ROM).
15. **Cui S., Goldsmith A., Bahai A.** Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks // IEEE JSAC. Aug. 2004. Vol. 22. P. 1089–1098.
16. **Jayaweera S.** An energy-efficient virtual MIMO architecture based on V-BLAST processing for distributed wireless sensor networks // Proc. IEEE SECON. Oct. 2004. P. 299–308.
17. **Stuber G. L.** Principles of Mobile Communication. — Norwell, MA: Kluwer Academic Publishers, 2001. — 776 p.
18. **Rabaey J., Ammer J., da Silva J. Jr., Patel D.** PicoRadio: Ad-hoc wireless networking of ubiquitous low-energy sensor/monitor nodes // Proc. IEEE VLSI Workshop, 2000. — 1 electron. opt. disk (CD-ROM).
19. **Haenggi M., Puccinelli D.** Routing in ad hoc networks: a case for long hops // IEEE Commun. Magazine. Oct. 2005. Vol. 43. P. 112–119.
20. **Rong B., Ephremides A.** On opportunistic cooperation for improving the stability region with multipacket reception // Lecture Notes in Computer Science. Springer, 2009. Vol. 5894/2009. P. 45–59.
21. **Jaiswal N.** Priority Queues. — N. Y.: Academic Press, 1968. — 240 p.
22. **Kleinrock L.** Queueing Systems. Vol. 1. Theory. — John Wiley & Sons, 1975. — 417 p.

УДК 004.05

УЛУЧШЕНИЕ СПОСОБОВ АУТЕНТИФИКАЦИИ ДЛЯ КАНАЛОВ СВЯЗИ С ОШИБКАМИ

В. Н. Никитин,
канд. техн. наук, доцент
Д. В. Юркин,
аспирант
Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича

Рассмотрен обобщенный подход к разработке и анализу криптографических протоколов с помощью вероятностно-временных методов. Показано влияние ошибок, возникающих в канале связи, на работу протоколов аутентификации.

Ключевые слова *— криптографический протокол, канал связи с ошибками, вероятностно-временные характеристики.*

Введение

Одним из основных показателей систем конфиденциальной связи, наряду со стойкостью и трудоемкостью выполнения, является эффективность использования ресурсов сети связи, обеспечивающей для корреспондентов-участников криптографического протокола своевременные предоставление доступа и передачу данных. Поэтому обеспечение высокого качества конфиденциальной связи невозможно без обеспечения заданных требований к вероятностно-временным характеристикам криптографических протоколов предоставления доступа к защищенному каналу связи и инкапсуляции данных. Наибольшей актуальностью обладает проблема идентификации и аутентификации корреспондентов в сетях широкополосного радиодоступа.

Задачи совершенствования способов аутентификации для каналов связи с ошибками

Исходя из общих требований к безопасности защищенных каналов связи можно сформулировать следующие требования к протоколам предоставления доступа [1], которые имеют в своей основе криптографические методы аутентификации:

1) сообщения протокола должны быть результатом выполнения однонаправленного преобразования, обусловленного знанием общего секрета на основе преобразования запросов и общего секрета, не компрометирующего общий секрет;

2) во избежание атаки повторения ранее переданных запросов должна быть обеспечена устойчивость к накоплению статистики передаваемых сообщений [2].

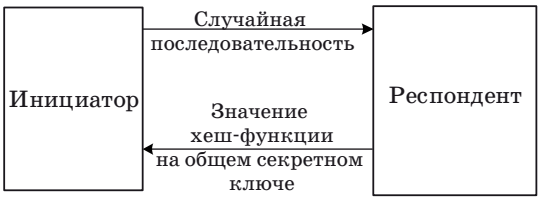
Вместе с тем, основываясь на требованиях по своевременности предоставления доступа в сетях передачи данных общего пользования [3], необходимо также обеспечить следующие вероятностно-временные требования:

- среднее время успешной аутентификации $\overline{T}$;
- вероятность успешного выполнения протокола аутентификации за допустимое время $\overline{P}(\overline{T} \le T_{exec})$.

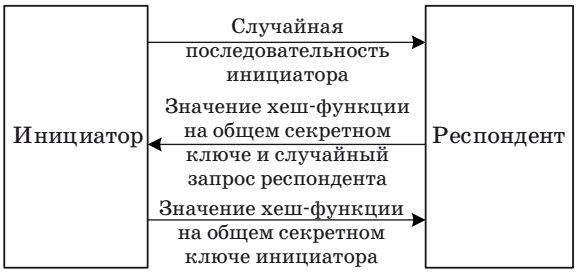
Таким образом, целью разработки является улучшение вероятностно-временных характеристик протокола аутентификации без снижения стойкости к атакам за счет сокращения числа передаваемых сообщений и уменьшения временных затрат на периодическую смену ключевой информации при применении протокола аутентификации в каналах связи с ошибками, а решаемые при этом задачи должны обеспечить достижение максимальной вероятности предоставления доступа за допустимое время в канале связи с ошибками.

Аутентификация ISO/IEC 9798 и RIPE-RACE

Известен [4–8] ряд способов аутентификации (рис. 1), использующих модель информационного



■ ***Рис. 1.*** *Способ односторонней аутентификации ISO/IEC 9798*



■ ***Рис. 2.*** *Способ двусторонней аутентификации RIPE-RACE*

взаимодействия корреспондентов типа «запрос-ответ».

В семействе стандартов ISO/IEC 9798 описан способ аутентификации, использующий модель «запрос-ответ» с применением ключевой хеш-функции, реализующий одностороннюю аутентификацию корреспондентов. Сущность способа заключается в аутентификации на основании формирования и передачи случайного запроса инициатором и вычисления респондентом ключевой хеш-функции от него с последующей передачей ее значения в ответ. Недостатком этого способа является отсутствие возможности двусторонней аутентификации корреспондентов за одну сессию протокола. Это обусловлено тем, что алгоритмы вычислений и информационного обмена сообщениями не позволяют выполнить случайные запросы инициатора и респондента в одной сессии протокола. При этом задача двусторонней аутентификации может быть решена двукратным выполнением [4] односторонней аутентификации, что потребует существенного увеличения временных затрат.

Решением задачи двусторонней аутентификации является способ RIPE-RACE [9] (рис. 2), в котором также используют ключевую хеш-функцию.

В данном способе инициатор передает случайный запрос респонденту, который, получив этот запрос, формирует свой случайный запрос, вычисляет ключевую хеш-функцию от обоих запросов и передает свой случайный запрос вместе с результатом вычисления ключевой хеш-функции. Инициатор, получив ответ респондента, вычисляет, используя секретный ключ, хеш-функцию от своего запроса, полученного запроса и идентификатора респондента и сравнивает полученное значение с принятым. В случае совпадения он вычисляет значение хеш-функции на том же ключе от обоих запросов и своего идентификатора, после чего передает его респонденту. Респондент, приняв сообщение инициатора, также вычисляет хеш-функцию от обоих запросов и сравнивает полученное значение с принятым. В случае совпадения сравниваемых величин протокол аутентификации завершен успешно.

Такой способ позволяет уменьшить число передаваемых сообщений с четырех до трех, что обеспечивает сокращение времени аутентификации для систем связи, работающих по различным каналам связи, не снижая вычислительной стойкости способа прототипа. Это повышает эффективность использования пропускной способности канала связи и снижает временные затраты на получение доступа к информационному ресурсу.

Однако рассмотренные способы имеют два недостатка:

— число передаваемых сообщений избыточно;

— криптосистемы, используемые в нем, требуют периодической смены общего секрета.

Аутентификация с использованием бесключевых хеш-функций

В рамках выбранной модели информационного взаимодействия для выполнения требуемого преобразования при аутентификации можно использовать любую условно однонаправленную функцию. Если в качестве такой функции выбрана ключевая хеш-функция [10], то заранее распределенная последовательность используется как общий секрет для вычисления однонаправленных преобразований от случайных запросов, и в целях защиты от статистических атак необходима постоянная смена общего секрета.

Однонаправленное преобразование можно реализовать с использованием другого класса функций, которым может являться класс бесключевых криптографических хеш-функций [11].

В предлагаемом способе двусторонней аутентификации корреспондентов системы связи и аутентификации при определении доступа субъекта к информационным ресурсам технических средств передачи хранения и обработки информации передается только два сообщения, вместо ключевых хеш-функций для преобразования информации используются бесключевые хеш-

функции, а общий секрет применяется в качестве аргумента хеш-функции.

Случайность однонаправленного преобразования достигается за счет конкатенации общего секрета со случайным аргументом хеш-функции, что позволяет при выборе стойкой однонаправленной функции такого класса приравнять вероятность успешной атаки на алгоритм, основанной на вычислении общего секрета, составляющего ее аргумент, к вероятности успешного обращения этой функции.

Протокол выполняется следующим образом (рис. 3). Первым сообщением передается запрос C инициатора и ответ h_R на заранее известный только легитимным корреспондентам запрос (общий секрет), а вторым — ответ h_S респондента инициатору. Причем запрос инициатора формируется путем вычисления бесключевой хеш-функции (например, алгоритмы [12, 13]) аргумента, составляющего результат вычисления $h_R = h(h_s \| C) = h(SAB \| h_s(SAB \| C)$ той же хеш-функции $h(x)$ случайного запроса и общего секрета $h_S = h(SAB \| C)$, конкатенированного со значением случайного числа, к которому добавлено само случайное число.

Ответное сообщение респондента состоит из значения бесключевой хеш-функции $h_s = h(S_{AB} \| C)$ от общего секрета и случайного запроса инициатора.

При таком информационном обмене (рис. 4) за счет уменьшения количества передаваемых сообщений сокращаются временные затраты на выполнение двусторонней аутентификации корреспондентов, а стойкость способа определяется выбором алгоритма вычисления однонаправленной бесключевой хеш-функции. Такой способ не требует периодической смены общего секрета, что позволяет сделать его долгосрочным.

Вследствие этого при работе по каналам связи с ошибками увеличивается вероятность успешного завершения протокола в заданное время, что позволяет сократить время доступа к защищенному каналу связи при заданной вероятности успешной аутентификации и улучшает доступность информационного ресурса для легитимных корреспондентов субъектов информационного обмена.

Сравнение вероятностно-временных характеристик протоколов аутентификации модели «запрос-ответ»

Приведем оценки среднего времени $\bar{T}(p_{oo})$ выполнения и вероятности успешного завершения $\bar{P}(p_{oo})$ от вероятности обнаружения ошибки в сообщениях протокола.

Согласно методике оценки вероятностно-временных характеристик криптографических протоколов [14] для рассматриваемых протоколов двусторонней аутентификации ISO/IEC 9798, SKID и предлагаемого протокола с использованием бесключевой хеш-функции при предоставлении доступа, имеем

$$\bar{T}(l, v_{form}, v_{form}, p_{oo}) = \left. \frac{\mathrm{d}}{\mathrm{d}x} \prod_{i=1}^{j} \frac{f_i^{form}(l, v_{form}, p_{oo}, x) f_i^{send}(l, v_{send}, p_{oo}, x)}{1 - f_i^{err}(l, v_{send}, p_{oo}, x) f_i^{form}(l, v_{form}, p_{oo}, x)} \right|_{x=1};$$

$$\bar{P}(l, p_{oo}, z, v_{form}, v_{send}) = 1 - \left(1 - (1 - p_{oo})^{\sum_{i=1}^{j} l_j}\right)^{\left. \mathrm{floor}\left(\frac{T_{exec}}{\bar{T}}\right) \right|_{p_{oo}=0}},$$

где l — длина кодируемого сообщения; v_{form}, v_{send} — скорости формирования и передачи сообщения соответственно; p_{oo} — вероятность обнаруженной битовой ошибки; j — число сообщений протокола.

На рис. 5, *а*, *б* показаны зависимости среднего времени и вероятности успешного завершения при $T_{exec} = 10$ с для сравниваемых протоколов применительно к сетям передачи данных стандарта IEEE 802.11 для следующих исходных данных:

скорость формирования кадра — $2 \cdot 10^9$ бит/с;

длина аргумента хеш-функции — 64 бит;

■ ***Рис. 3.*** *Информационное взаимодействие корреспондентов при двусторонней аутентификации*

Предварительные вычисления				
Распределение общего секрета				
Инициатор			Респондент	
Вычисления	Прием / передача		Прием / передача	Вычисления
$rnd(C)$ $h_S = h(S_{AB}\|C)$ $h_R = h(h_s\|C) = h(S_{AB}\|h_s(S_{AB}\|C))$				
Активные вычисления				
	C h_R	→	C h_R	$\widetilde{h}_S = \widetilde{h}(S_{AB}\|C)$ $\widetilde{h}_R = h(h_s\|C) = h(S_{AB}\|h_s(S_{AB}\|C))$ $\widetilde{h}_R = h_R$
$h_S = \widetilde{h}_S$	$\widetilde{h}_S$	←	$\widetilde{h}_S$	

■ ***Рис. 4.*** *Схема выполнения корреспондентами протокола двусторонней аутентификации*

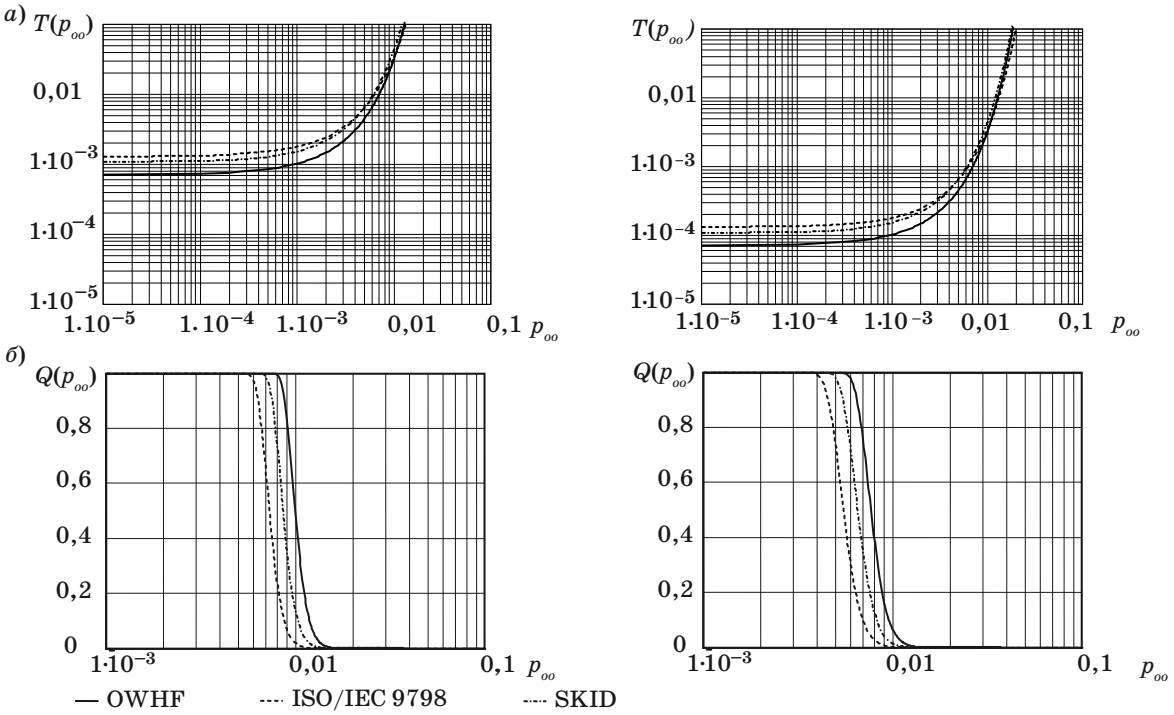
длина значения хеш-функции — 128 бит;

количество служебной информации на кадр — 224 бит.

По результатам оценки вероятностно-временных характеристик рассмотренных протоколов необходимо отметить значительное улучшение среднего времени выполнения и вероятности успешного завершения в заданное время предлагаемого способа по отношению к аналогам. Данное преимущество характеризует протокол аутентификации с использованием алгоритмов вычисления бесключевых хеш-функций как наилучший по производительности способ предоставления доступа субъекта к информационному ресурсу, эффективно работающий по каналам с высокой вероятностью ошибки.

Заключение

Анализ известных способов аутентификации модели «запрос-ответ» показывает, что они облада-



■ ***Рис. 5.*** *Зависимости среднего времени выполнения (а) и вероятности успешного завершения (б) протоколов двусторонней аутентификации: скорость передачи кадра Frame Management 10^6 бит/с (слева) и 10^7 бит/с (справа)*

ют недостатками, затрудняющими их использование в каналах низкого качества. В работе предложен способ аутентификации, позволяющий уменьшить время выполнения протокола двусторонней аутентификации, что достигается за счет сокращения числа раундов передачи сообщений до минимально возможного. Это обеспечивается вычислением бесключевой хеш-функции над конкатенацией общего секрета со случайной величиной, изменяемой при выполнении каждой последующей итерации протокола, что посредством рандомизации обеспечивает защиту от накопления статистики и атаки повторных передач сообщений. Таким образом, уникальность результата выполнения однонаправленного преобразования общего секрета достигается посредством рандомизации запроса путем добавления к аргументу вычисляемой хеш-функции случайной последовательности.

Предлагаемый способ аутентификации с использованием бесключевых хеш-функций позволяет существенно сократить время предоставления доступа в радиоканалах низкого качества.

Литература

1. **U.S.** Department of Commerce/National Bureau of Standards. Password usage. National Technical Information Service. — Virginia: Springfield, 1985. www.itl.nist.gov/fipspubs/fip112.htm (дата обращения: 10.11.08).
2. **Gong L.** Variations on the themes of message freshness and replay// The Computer Security Foundations Workshop. Geneva: IEEE Computer Society Press, 1993. P. 131–136.
3. **РД 45.128-2000.** Сети и службы передачи данных/ Министерство Российской Федерации по связи и информатизации, 2001. http://minkomsvjaz.ru/ministry/documents/959/ (дата обращения: 10.11.08).
4. **ISO/IEC 9798-1.** Information technology — Security techniques — Entity authentication mechanisms. Part 1: General model/International Organization for Standardization. — Geneva, 1991. http://www.iso.org/iso/iso_catalogue/catalogue_tc/ (дата обращения: 10.11.08).
5. **SO/IEC 9798-4.** Information technology — Security techniques — Entity authentication. Part 4: Mechanisms using a cryptographic check function/International Organization for Standardization. — Geneva, 1995. http://www.iso.org/iso/iso_catalogue/catalogue_tc/ (дата обращения: 10.11.08).
6. **ISO/IEC 9798-2.** Information technology — Security techniques — Entity authentication. Part 2: Mechanisms using symmetric encipherment algorithms/International Organization for Standardization. — Geneva, 1994. http://www.iso.org/iso/iso_catalogue/catalogue_tc/ (дата обращения: 10.11.08).
7. **ISO/IEC 9798-3.** Information technology — Security techniques — Entity authentication mechanisms. Part 3: Entity authentication using a public-key algorithm/ International Organization for Standardization. — Geneva, 1993. http://www.iso.org/iso/iso_catalogue/catalogue_tc/ (дата обращения: 10.11.08).
8. **ISO/IEC 9798-5.** Information technology — Security techniques — Entity authentication. Part 5: Mechanisms using zero knowledge techniques/International Organization for Standardization. — Geneva, 1996. http://www.iso.org/iso/iso_catalogue/catalogue_tc/ (дата обращения: 10.11.08).
9. **Bosselares A., Preneel B.** Integrity Primitives for Secure Information Systems//Final Report of RACE Integrity Primitives Evaluation RIPE-RACE 1040. N. Y.: Springer-Verlag, 1995. P. 1–12.
10. **Metzger P., Simpson W.** IP authentication using keyed MD5. 1995. http://www.ietf.org/rfc/rfc1852.txt (дата обращения: 10.11.08).
11. **Preneel B., Leuven K. U., Mercierlaan K.** Cryptographic Hash Functions: an overview// ESAT-COSIC Laboratory. Leuven, Belgium, 1994. P. 412–431.
12. **Matyas S. M., Meyer C. H., Oseas J.** Generating strong one-way functions with cryptographic algorithm// IBM Technical Disclosure Bulletin. Mar. 1985. Vol. 27. N 10A. P. 5658–5659.
13. **Miyaguchi S., Ohta K., Iwata M.** 128-bit hash function (N-hash)//NTT Review. Nov. 1990. Vol. 2. N 6. P. 128–132.
14. **Nikitin V., Yurkin D., Chilamkurti N.** The influence of the cryptographic protocols on the quality of the radio transmission// ICUMT. St.-Petersburg, Russia, Nov. 2009. P. 1–5.

УДК 519.688

КОДОВЫЕ ШАРАДЫ

А. Л. Чмора,
ведущий специалист
ОАО «Инфотекс»

Рассматривается метод противодействия DoS-атаке с использованием шарад, построенных на основе кодов, исправляющих ошибки. Обосновывается практическая состоятельность итеративного метода конструирования шарад. Итеративные кодовые шарады позволяют исключить применение квантового компьютера и массированного распараллеливания в целях снижения трудоемкости отыскания решения.

Ключевые слова *— DoS-атака, вычислительные шарады, коды, исправляющие ошибки, линейные коды.*

Введение

DoS-атаки (*Denial of Service*) широко распространены в сети Интернет [1, 2]. Задача атакующего – создать искусственную ситуацию, в которой добросовестному потребителю будет отказано в предоставлении соответствующих услуг. Для объяснения воспользуемся следующей бытовой аналогией. Предположим, что в ресторане имеется некоторое количество столиков и каждый можно зарезервировать, позвонив по телефону. Звонки принимаются до часа дня включительно. Злоумышленник, воспользовавшись простейшей стратегией, способен причинить ресторану ощутимый финансовый и репутационный ущерб. Для этого достаточно зарезервировать все доступные столики до установленного часа. Очевидно, что если все столики уже зарезервированы, то большинство потенциальных посетителей откажется от запланированного посещения и предпочтет другой ресторан. Конечно же, в какой-то момент руководство ресторана обнаружит факт злоупотребления и аннулирует резервирование, но с непренебрежимой вероятностью некоторое количество столиков в течение вечера не будет востребовано.

Сетевая DoS-атака

Поглощающая ресурсы стратегия относится к наиболее распространенному типу DoS-атаки. Так, злоумышленник способен инициировать аномальное количество сетевых соединений, но предусмотренные протоколом правила взаимодействия при этом умышленно игнорируются. Рассмотрим подробнее, как это происходит на примере TCP-соединения [3].

Обычно TCP-соединение устанавливается при помощи полуторараундного протокола [1]. Клиент инициирует соединение, передав серверу специальный SYN-пакет. В ответ сервер передает пакет SYN ACK и тем самым подтверждает соединение. По факту подтверждения соединения сервер выделяет некоторую область памяти. Подчеркнем, что объем выделенной под соединения памяти конечен. В завершение клиент выполняет квитирование и передает серверу ACK-пакет. Атакующий инициирует множество соединений, в которых пакеты SYN и SYN ACK передаются в соответствии с протоколом, но умышленно опускает квитирование, т. е. ACK-пакет никогда не передается. В итоге соединение не устанавливается, но при этом сервер не освобождает выделенную область памяти. Если количество таких незавершенных соединений достаточно велико, то происходит переполнение памяти и сервер перестает реагировать на какие-либо запросы. Понятно, что на прикладном уровне описанная сетевая атака приводит к отказу в обслуживании.

Вычислительные шарады

Воспользуемся вычислительной задачей, для которой трудоемкость отыскания решения варьируется в широком диапазоне значений и задается параметрически. Для этой цели подходят такие задачи, про которые известно, что для них не существует иных, более эффективных, методов

[1] Раунд состоит из запроса и ответного подтверждения. Полуторараундный протокол включает дополнительное сообщение-квитанцию в ответ на подтверждение.

решения. Например, никакие выполненные заранее предвычисления не способствуют снижению трудоемкости. Решение может быть найдено исключительно при помощи *силовой атаки*, т. е. методом проб и ошибок с исчерпывающим перебором вариантов. Для обозначения таких задач воспользуемся термином *шарада* (puzzle). Идеологическая подоплека метода шарад, а также исследование их свойств восходят к пионерской работе Р. Меркля [4].

Назовем *экзаменатором* того, кто создает шараду и по построению располагает ее решением, а *экзаменуемым* того, кто выполняет отыскание решения по заданию экзаменатора.

В работе [5] предложен практический метод противодействия сетевой атаке. Сервер предлагает решить шараду каждый раз, когда устанавливается соединение. Память под соединение выделяется только при условии предоставления правильного решения. Атакующий продуцирует запросы на установление соединения. Поскольку число таких запросов аномально велико, и в этом суть атаки, то и число шарад также велико. Искусственно созданная сетевая нагрузка возвращается к атакующему в виде вычислительной нагрузки, и для достижения поставленной цели он вынужден инвестировать. Тогда DoS-атака перестает быть беззатратной и атакующий вынужден платить за ее осуществление. Добросовестный пользователь, в отличие от атакующего, продуцирует умеренное число запросов, и для него вычислительная нагрузка не обременительна. Следует подчеркнуть, что эффективность механизма противодействия обусловлена исключительно разницей в количестве запросов на установление соединения. В работах [6—8] метод шарад применяется для противодействия DoS-атаке в ряде других приложений.

Сформулируем набор требований к шарадам в контексте DoS-атаки.

1. Собственно шарады не должны быть инструментом атаки. Вычислительная трудоемкость построения шарады и проверки ее решения не должна быть чрезмерной.

2. Трудоемкость решения шарады должна быть регулируемой. Необходимо, чтобы сервер имел возможность гибко настраивать трудоемкость, оперативно реагируя на увеличение или снижение сетевой нагрузки.

3. Решение шарады возможно при наличии определенного вычислительного потенциала. Алгоритм решения должен быть задан строго. Трудоемкость решения должна быть ограничена сверху. Решение может быть получено за конечное время. Не должно существовать известных методов повышения эффективности решения.

Недостатки метода шарад

Известные шарады [5, 9] допускают возможность отыскания решения независимыми вычислителями, причем каждый из таких вычислителей выполняет поиск в пределах некоторого подмножества претендентов, мощность которого меньше мощности исходного множества. Назовем такой подход к поиску решения *распараллеливанием*. На практике это означает, что атакующий может воспользоваться методом распределенных вычислений. Тогда применение N независимых вычислителей или вычислителя с N-ядерной архитектурой позволяет получить результат в N раз быстрее. Очевидно, что для организации таких вычислений атакующему необходимо выполнить предварительную подготовку заданий с последующим их распределением при помощи специализированного протокола. Как только решение найдено одним из вычислителей, все остальные должны по команде прекратить обработку заданий.

Напротив, шарады с последовательным алгоритмом решения [10] не допускают распараллеливания, и в этом их бесспорное преимущество. Однако эти шарады уязвимы с точки зрения атаки при помощи квантового компьютера. Известно, что трудоемкость отыскания решения шарады [10] может быть эффективно снижена в результате факторизации. Метод с полиномиальной трудоемкостью, известный как алгоритм факторизации Шора [11, 12], в существенной мере использует принцип квантовых вычислений. Объявлено о создании прототипа программируемого квантового компьютера с ограниченными вычислительными возможностями [13]. Известен также пример факторизации при помощи алгоритма Шора на реальном квантовом вычислителе [14]. Для эффективной факторизации числа из k двоичных разрядов понадобится квантовый вычислитель с регистром на $K \approx 2k$ квантовых состояний (кубит) [15]. Логично ожидать появления полноценного квантового компьютера в ближайшие десятилетия. Шарады [9] не только эффективно решаются при помощи квантового вычислителя [11], но также допускают распараллеливание.

Выделим следующие типы шарад с регулируемой трудоемкостью решения.

1. Шарады, допускающие распараллеливание с применением квантовых/неквантовых вычислителей, для которых неизвестен эффективный квантовый алгоритм решения [5].

2. Шарады, допускающие распараллеливание с применением квантовых/неквантовых вычислителей, для которых известен эффективный квантовый алгоритм решения [9].

3. Шарады, не допускающие распараллеливания с применением квантовых/неквантовых вычислителей, для которых известен эффективный квантовый алгоритм решения [10].

Таким образом, к недостаткам шарад перечисленных типов следует отнести **возможность распараллеливания и существование эффективного квантового алгоритма решения**.

Следует также упомянуть отсутствие строгого доказательства того факта, что фундаментальные задачи, лежащие в основе известных шарад, относятся к классу *NP*-трудных. По мнению ряда специалистов, наличие подобного доказательства предоставляет достаточные гарантии относительной неуязвимости в долгосрочной перспективе.

Постановка задачи

Задача заключается в конструировании шарады с регулируемой трудоемкостью отыскания решения, максимально широким диапазоном трудоемкости с возможностью плавной регулировки, минимальными объемом памяти и накладными расходами при передаче, которая не поддается эффективному распараллеливанию и для которой не известен эффективный квантовый алгоритм. Для решения задачи воспользуемся методами теории помехоустойчивого кодирования и линейной алгебры.

Кодовые шарады

Мотивация подхода в том, что вопрос об эффективном решении на квантовом компьютере фундаментальных задач теории помехоустойчивого кодирования в настоящее время остается открытым. Более того, в работе [16] показано, что эти задачи относятся к классу *NP*-трудных. Известно также [17], что трудоемкость зашифрования для кодовой асимметричной криптосистемы ниже, чем трудоемкость зашифрования для криптосистемы RSA. Логично предположить, что шарады на основе кодов не превосходят по трудоемкости построения шарады из работы [10] и сравнимы с шарадами из работы [9].

Пусть имеется k-мерный линейный код C с минимальным расстоянием d [18]. Код задан $k\times n$ порождающей матрицей $\mathbf{G}$. Тогда существует кодовое слово $\mathbf{c}=\mathbf{pG}$, $\mathbf{c}\in\ker(\mathbf{H})$, где $\mathbf{H}$ — $(n-k)\times n$ проверочная матрица кода C и $\mathbf{p}$ — информационная последовательность. Множество $\mathbb{F}_q^n=\{\mathbf{x}=(x_1, \ldots, x_n)\,|\,x_j\in\mathbb{F}_q\}$ рассматривается как линейное пространство размерности n над $\mathbb{F}_q$.

Через $h(\cdot)$ обозначим криптографическую хэш-функцию $f_\lambda:\{0,1\}^*\to\{0,1\}^\lambda$ для некоторого λ. Свойства $h(\cdot)$ подробно описаны [19]. Отметим, что $h(\cdot)$ доступна как экзаменатору, так и экзаменуемому.

Назовем *кодовыми* шарады, построенные на основе кода C. Отметим, что код C известен как экзаменатору, так и экзаменуемому.

Для заданного $[n, k, d]_q$-кода экзаменатор выполняет следующие действия.

1. Выбирает информационную последовательность $\mathbf{p}\in_R\mathbb{F}_q^k$.
2. Сохраняет $\varphi=h(\mathbf{p})$ для проверки решения.
3. Выбирает ошибку $\mathbf{e}\in_R\mathbb{F}_q^n$ такую, что вес Хэмминга $wt(\mathbf{e})=t$ и $\lceil(d-1)/2\rceil<t<\lceil n/2\rceil$.
4. Вычисляет значение хэш-функции $\psi=h(\mathbf{p}\,\|\,\mathbf{e})$.
5. Вычисляет сумму $\check{\mathbf{c}}=\mathbf{pG}+\mathbf{e}$.
6. Передает $\{\check{\mathbf{c}}, \psi, t\}$ экзаменуемому.

Для поиска решения экзаменуемый выполняет следующие действия.

1. Выбирает ошибку $\mathbf{e}\in\mathbb{F}_q^n$ такую, что $wt(\mathbf{e})\le t$.
2. Вычисляет сумму $\hat{\mathbf{c}}=\check{\mathbf{c}}+\mathbf{e}$.
3. Выполняет декодирование $\hat{\mathbf{c}}$ и получает $\hat{\mathbf{p}}$.
4. Проверяет $h(\hat{\mathbf{p}}\,\|\,\mathbf{e})\stackrel{?}{=}\psi$.
5. Если $h(\hat{\mathbf{p}}\,\|\,\mathbf{e})=\psi$, то переходит к 6, иначе к 1.
6. Предъявляет $\varphi'=h(\hat{\mathbf{p}})$ в качестве решения.

Заметим, что в кодовой шараде секретный ключ состоит из двух компонент – информационной последовательности $\mathbf{p}$ и ошибки $\mathbf{e}$. Если известно $\mathbf{p}$, то легко вычислить $\mathbf{e}$, так как $\mathbf{e}=\check{\mathbf{c}}+\mathbf{pG}$, и наоборот.

Вес t всегда должен превышать половину кодового расстояния на величину ε, которая обеспечивает маскировку кода с алгоритмом декодирования полиномиальной трудоемкости под код, для которого возможно только корреляционное декодирование. Размерность кода следует выбирать так, чтобы трудоемкость поиска решения перебором по $\mathbf{p}$ превышала трудоемкость поиска решения перебором по $\mathbf{e}$. Действительно, если

$$q^k<\sum_{i=\lceil(d-1)/2\rceil+\varepsilon}^{t}(q-1)^i\binom{n}{i}, \qquad (1)$$

то экзаменуемый найдет решение шарады $\{\check{\mathbf{c}}, \psi, t\}$ в среднем за q^{k-1} испытаний, воспользовавшись следующим методом.

1. Выбирает информационную последовательность $\hat{\mathbf{p}}\in\mathbb{F}_q^k$.
2. Вычисляет кодовое слово $\hat{\mathbf{c}}=\hat{\mathbf{p}}\mathbf{G}$.
3. Вычисляет сумму $\mathbf{e}=\hat{\mathbf{c}}+\check{\mathbf{c}}$.
4. Проверяет $h(\hat{\mathbf{p}}\,\|\,\mathbf{e})\stackrel{?}{=}\psi$.
5. Если $h(\hat{\mathbf{p}}\,\|\,\mathbf{e})=\psi$, то переходит к 6, иначе к 1.
6. Предъявляет $\varphi'=h(\hat{\mathbf{p}})$ в качестве решения.

Использование критерия $\psi=h(\mathbf{p}\,\|\,\mathbf{e})$ не позволяет экзаменуемому получить искомое решение за меньшее число испытаний. Предположим, $\psi=h(\mathbf{p})$ и необходимо найти $\mathbf{p}$. Пусть выбрано $\check{\mathbf{e}}$

такое, что $wt(\check{\mathbf{e}}+\mathbf{e}) \le \lfloor (d-1)/2 \rfloor$. Если вес ошибки не превышает половины кодового расстояния, то в результате декодирования последовательность $\mathbf{p}$ будет восстановлена правильно и решение будет получено при $\check{\mathbf{e}} \ne \mathbf{e}$. Очевидно, что $h(\mathbf{p} \| \check{\mathbf{e}}) = \psi$ только при $\check{\mathbf{e}} = \mathbf{e}$.

Кодовая шарада должна допускать регулировку трудоемкости отыскания решения. Вес Хэмминга t ошибки $\mathbf{e}$ — один из параметров, определяющий объем перебора при поиске решения. Обозначим через $\mathcal{R}$ множество всевозможных претендентов. Известно, что $|\mathcal{R}|$ достигает своего максимума при $t = \lfloor n/2 \rfloor$. Например, для $q = 2$ существует $\sum_{i=\lfloor (d-1)/2 \rfloor + \varepsilon}^{t} \binom{n}{i} \le \frac{n}{(n/2-t)^2} 2^{n-3}$ претендентов на $\mathbf{e}$. Для поиска решения при $t = n/2 - 1$ экзаменуемому потребуется выполнить в среднем не более $n2^{n-4}$ испытаний. С другой стороны, следует выбирать параметры кода так, чтобы трудоемкость отыскания решения варьировалась в широком диапазоне. Из указанных ограничений следует, что $\lfloor (d-1)/2 \rfloor + \varepsilon \le t \le \lfloor n/2 \rfloor$ и

$$\lfloor (d-1)/2 \rfloor + \varepsilon \le \lfloor n/2 \rfloor - 1. \qquad (2)$$

Например, если предположить равенство в (2), то при фиксированном n существует единственная шарада с максимальной трудоемкостью.

Назовем шараду *устойчивой*, если не существует иного, менее трудоемкого способа ее решения, кроме заданного по построению. Необходимо определить допустимый диапазон скоростей кода такой, чтобы гарантировать устойчивость кодовой шарады и обеспечить максимально широкий диапазон трудоемкости.

Чем меньше минимальное кодовое расстояние d, тем шире диапазон трудоемкости. Очевидно, что d обратно пропорционально размерности кода. Это означает, что для конструирования шарад предпочтительнее высокоскоростные коды, для которых отношение $R = k/n$ стремится к 1.

Пусть задан $[n, n-d+1, d]_q$-код Рида—Соломона $RS_q(n, d)$ над $\mathbb{F}_q$, $q = p^m$, где p — простое число; m — положительное целое, который имеет максимально возможную размерность при заданных n и d [18]. Тогда $d = n-k+1$ и код исправляет $t \le \lfloor (n-k)/2 \rfloor$ ошибок. Существуют $RS_q(n, d)$-коды с блоковой длиной $n = q-1$, расширенные с $n = q$ и дважды расширенные с $n = q+1$. Это означает, что всегда можно выбрать код с четным n. Шараду на основе $RS_q(n, d)$-кода будем называть $RS_q(n, d)$-шарадой.

Не существует обоснованных возражений против использования безызбыточных кодов. Тогда $RS_q(n, 1)$-шарада обладает максимально широким диапазоном трудоемкости, поскольку $\varepsilon = 0$ и вес ошибки варьируется в интервале $n/2 \ge t \ge 1$. Шарада безусловно устойчива, так как $k > n/2$. Следовательно, верхняя граница скорости кода совпадает с конструктивным ограничением.

Нижняя граница может быть получена при помощи следующих простых рассуждений. Очевидно, что при $k \le n/2$ наблюдается объективное сужение диапазона трудоемкости, поскольку вес ошибки $\lfloor (d-1)/2 \rfloor + \varepsilon \le t < k$. Если $k < t \le n/2$, то шарада неустойчива. Это объясняется тем, что справедливо неравенство (1) и трудоемкость отыскания решения будет ниже запланированной. Таким образом, для построения устойчивых $RS_q(n, d)$-шарад с широким диапазоном трудоемкости следует использовать коды со скоростями в интервале $0{,}5 < R \le 1$. $RS_q(n, 1)$-шарады представляются наиболее перспективными с практической точки зрения.

Отметим, что для $RS_q(n, 1)$-шарад ограничение на блоковую длину не является критичным, поскольку трудоемкость поиска решения в существенной степени определяется весом t.

$RS_q(n, d)$-шарады допускают распараллеливание. Попытаемся устранить этот недостаток. Воспользуемся следующим наблюдением. Как было отмечено, для организации параллельных вычислений необходимо выполнить распределение заданий. Если такое распределение выполняется однократно, то возникающими накладными расходами можно пренебречь. Предположим, что шарада состоит из нескольких подшарад. Процесс распределения заданий усложнится, если скомбинировать подшарады таким образом, что решение каждой последующей будет зависеть от решения предыдущей. Иначе говоря, атакующий будет вынужден распределять задания для каждой подшарады. При этом важно, чтобы подшарады не были известны заранее и раскрывались последовательно по мере получения промежуточных решений.

Сконструируем $RS_q^{\ell}(n, 1)$-шараду, состоящую из ℓ вложенных подшарад. Для этого зададим вес ошибки t_j для каждой из ℓ подшарад. В результате получим набор $\{t_1, \ldots, t_\ell\}$. Напомним, что для $RS_q(n, 1)$-шарады $t \in [1, n/2]$.

Для построения $RS_q^{\ell}(n, 1)$-шарады экзаменатор выполняет следующие действия.

1. Выбирает информационную последовательность $\mathbf{p} \in {}_R\mathbb{F}_q^n$.

2. Сохраняет $\varphi = h(\mathbf{p})$ для проверки решения.

3. Устанавливает $j := 1$ и $\check{\mathbf{p}} := \mathbf{p}$.

4. Выбирает ошибку $\mathbf{e} \in {}_R\mathbb{F}_q^n$ такую, что $1 \le wt(\mathbf{e}) \le t_j$.

5. Вычисляет значение хэш-функции $\psi_j = h(\check{\mathbf{p}} \| \mathbf{e})$.

6. Вычисляет $\check{\mathbf{c}} = \check{\mathbf{p}}\mathbf{G} + \mathbf{e}$.

7. Устанавливает $j := j+1$ и $\check{\mathbf{p}} := \check{\mathbf{c}}$.

8. Проверяет $j \stackrel{?}{=} \ell + 1$.

9. Если $j = \ell + 1$, то переходит к 10, иначе к 4.

10. Передает $\{\check{\mathbf{c}}, \ell, \{\psi_1, ..., \psi_\ell\}, \{t_1, ..., t_\ell\}\}$ экзаменуемому.

Для поиска решения экзаменуемый выполняет следующие действия.

1. Устанавливает $j := \ell$ и $\mathbf{p} := \check{\mathbf{c}}$.

2. Выбирает ошибку $\mathbf{e} \in \mathbb{F}_q^n$ такую, что $1 \le wt(\mathbf{e}) \le t_j$.

3. Вычисляет сумму $\tilde{\mathbf{c}}$.

4. В результате декодирования $\tilde{\mathbf{c}}$ получает $\check{\mathbf{p}}$.

5. Проверяет $h(\check{\mathbf{p}} \| \mathbf{e}) \stackrel{?}{=} \psi_j$.

6. Если $h(\check{\mathbf{p}} \| \mathbf{e}) = \psi_j$, то переходит к 7, иначе к 2.

7. Устанавливает $j := j - 1$ и $\mathbf{p} := \check{\mathbf{p}}$.

8. Проверяет $j \stackrel{?}{=} 0$.

9. Если $j = 0$, то переходит к 10, иначе к 2.

10. Предъявляет $\varphi' = h(\check{\mathbf{p}})$ в качестве решения.

Конструкция $RS_q^\ell(n, 1)$-шарады такова, что объем памяти для хранения результирующей шарады равен объему памяти для хранения отдельной подшарады. Однако потребуется дополнительная память для хранения ℓ пар $\{\psi_j, t_j\}$. Заметим, что все ψ_j имеют одинаковую разрядность λ и $t_j \le n/2$. Тогда для хранения $RS_q^\ell(n, 1)$-шарады при $n = p^m$ необходимо зарезервировать $O\left(nm\log_2 p + \ell\left(\frac{nm\log_2 p}{2} + \lambda\right)\right)$ двоичных разрядов.

Из представленной конструкции следует, что применение безызбыточного кода вполне оправдано — объем памяти для хранения подшарад не зависит от ℓ. Однако для ψ_j и t_j объем памяти растет линейно по ℓ. С точки зрения резервирования памяти вклад ψ_j и t_j неравнозначен. Как правило, веса ошибок подчиняются монотонной зависимости. Следовательно, можно хранить не веса ошибок, а описание функции, при помощи которой несложно вычислить произвольное t_j. Объем памяти для хранения такого описания не зависит от ℓ. Будем считать, что накладными расходами, связанными с хранением t_j, можно пренебречь. Для хранения ψ_j, в противоположность t_j, может потребоваться относительно большой объем памяти. Например, если воспользоваться хэш-функцией SHA-256, то при $\ell = 100$ и $n = 2^8$ объем памяти для хранения всех ψ_j на порядок превысит объем памяти для хранения $\check{\mathbf{c}}$ и t_j, $1 \le j \le \ell$.

Рассмотрим такую конструкцию шарады, у которой объем памяти для хранения ψ_j не зависит от ℓ.

Назовем *итеративным хэшированием* преобразование вида $\psi_{\ell-1} = h(\underbrace{h(...h(h(s))...)}_{\ell \text{ раз}})$, где ℓ — число итераций. Финальное значение $\psi_{\ell-1}$ получается из стартового s.

Понадобится также следующее свойство кода C. Пусть $\mathbf{c}_1, \mathbf{c}_2 \in C$. Тогда $\mathbf{c}_3 = \mathbf{c}_1 + \mathbf{c}_2 = (\mathbf{p}_1 + \mathbf{p}_2)\mathbf{G}$ и $\mathbf{c}_3 \in C$.

Подойдем к конструированию шарады следующим образом. Вначале зададим вес ошибки t_j для каждой из ℓ подшарад. Затем для $s \in {}_R\mathbb{Z}$ методом итеративного хэширования вычислим $\psi_{\ell-1}$, $\psi_{\ell-2}$, ..., ψ_1, ψ_0. Отобразим каждый ψ_j на линейное пространство размерности n над $\mathbb{F}_q$. Предположим, $n = p^m$, $b = \left\lceil \frac{\lambda}{m\log_2 p} \right\rceil$, $1 \le b \le n$ и каждый ψ_j состоит из b q-ичных символов. В результате отображения получим $\Psi_j = \mathbf{0}_q^{n-b} \| \psi_j$, где $\mathbf{0}_q^{n-b}$ — последовательность из $n-b$ нулевых символов поля $\mathbb{F}_q$ и $\Psi_j \in \mathbb{F}_q^n$, $0 \le j \le \ell - 1$.

Заданы наборы $\{\Psi_{\ell-1}, ..., \Psi_1, \Psi_0\}$ и $\{t_1, ..., t_\ell\}$. Для построения $RS_q^\ell(n, 1)^*$-шарады экзаменатор выполняет следующие действия.

1. Выбирает информационную последовательность $\mathbf{p} \in {}_R\mathbb{F}_q^n$.

2. Сохраняет $\varphi = h(\mathbf{p})$ для проверки решения.

3. Устанавливает $j := 1$ и $\check{\mathbf{p}} := \mathbf{p}$.

4. Выбирает ошибку $\mathbf{e} \in {}_R\mathbb{F}_q^n$ такую, что $1 \le wt(\mathbf{e}) \le t_j$.

5. Вычисляет $\check{\mathbf{c}} = (\check{\mathbf{p}} + \Psi_{\ell-j})\mathbf{G} + \mathbf{e}$.

6. Устанавливает $j := j + 1$ и $\check{\mathbf{p}} := \check{\mathbf{c}}$.

7. Проверяет $j \stackrel{?}{=} \ell + 1$.

8. Если $j = \ell + 1$, то переходит к 9, иначе к 4.

9. Передает $\{\check{\mathbf{c}}, \ell, s, \{t_1, ..., t_\ell\}\}$ экзаменуемому.

Экзаменуемый выполняет следующие действия.

1. Устанавливает $j := \ell$, $\mathbf{p} := \check{\mathbf{c}}$ и $\hbar := h(s)$.

2. Выбирает ошибку $\mathbf{e} \in \mathbb{F}_q^n$ такую, что $1 \le wt(\mathbf{e}) \le t_j$.

3. Вычисляет сумму $\tilde{\mathbf{c}} = \mathbf{p} + \mathbf{e}$.

4. В результате декодирования $\tilde{\mathbf{c}}$ получает $\check{\mathbf{p}}$.

5. Отображает $\Psi_{\ell-j} = \mathbf{0}_q^{n-b} \| \hbar$.

6. Проверяет $(\check{\mathbf{p}} + \Psi_{\ell-j})\mathbf{G} \stackrel{?}{=} \tilde{\mathbf{c}} + \Psi_{\ell-j}\mathbf{G}$.

7. Если $(\check{\mathbf{p}} + \Psi_{\ell-j})\mathbf{G} = \tilde{\mathbf{c}} + \Psi_{\ell-j}\mathbf{G}$, то переходит к 8, иначе к 2.

8. Устанавливает $j := j - 1$, $\mathbf{p} := \check{\mathbf{p}} + \Psi_{\ell-j}$ и $\hbar := h(\hbar)$.

9. Проверяет $j \stackrel{?}{=} 0$.

10. Если $j = 0$, то переходит к 11, иначе к 2.

11. Предъявляет $\varphi' = h(\mathbf{p})$ в качестве решения.

Предположим, что для представления числа s в памяти достаточно λ двоичных разрядов. Тогда для хранения $RS_q^\ell(n, 1)^*$-шарады потребуется зарезервировать не более $O\left(nm\log_2 p + \frac{\ell nm\log_2 p}{2} + \lambda\right)$ двоичных разрядов.

Трудоемкость отыскания решения не превышает

$$\sum_{j=1}^{\ell}\sum_{i=1}^{t_j}(q-1)^i\binom{n}{i} \qquad (3)$$

испытаний.

Проанализируем $RS_q^{\ell}(n, 1)^*$-шараду на устойчивость. Соответствующее итеративное преобразование можно представить в виде

$$\check{\mathbf{c}} = ((((...((((\mathbf{p}+\Psi_{\ell-1})\mathbf{G}+\mathbf{e}_1)+\Psi_{\ell-2})\mathbf{G}+\mathbf{e}_2)+ \\ +...+\Psi_1)\mathbf{G}+\mathbf{e}_{\ell-1})+\Psi_0)\mathbf{G}+\mathbf{e}_{\ell}),$$

где $\mathbf{p}, \mathbf{e}_j \in {}_R\mathbb{F}_q^n$ и $\Psi_j \neq \Psi_i$ для $i \neq j$, $0 \le i, j \le \ell$.

Необходимо ответить на следующий вопрос: способен ли экзаменуемый, располагая набором $\{\check{\mathbf{c}}, \ell, s, \{t_1, ..., t_\ell\}\}$, а также последовательностью $\Psi_{\ell-1}$, ..., Ψ_1, Ψ_0, уменьшить запланированную трудоемкость поиска решения?

Конструкция $RS_q^{\ell}(n, 1)^*$-шарады напоминает луковицу. Доступ к некоторому внутреннему слою возможен только после удаления всех объемлющих слоев. Каждый слой ассоциирован с отдельной подшарадой. Слой с номером j считается удаленным, если найдено решение для j-й подшарады. Поиск решения начинается с первого внешнего слоя, который всегда доступен. Собственно решение шарады располагается в сердцевине луковицы — в последнем внутреннем слое.

Если воспользоваться геометрической интерпретацией, то каждое кодовое слово $RS_q(n, 1)$-кода располагается в центре сферы нулевого радиуса и все такие сферы не пересекаются. Число сфер равно числу кодовых слов, которое для $RS_q(n, 1)$-кода совпадает с мощностью $\mathbb{F}_q^n$. Тогда произвольная ошибка веса $0 < t \le n$ переводит кодовое слово $RS_q(n, 1)$-кода в другое кодовое слово того же кода с единичной вероятностью. Это означает, что каждая подшарада $RS_q^{\ell}(n, 1)^*$-шарады имеет единственное решение.

При заданном s несложно вычислить $\Psi_{\ell-j}$, которое используется в качестве критерия. Если $\mathbf{c} = (\mathbf{p}+\Psi_{\ell-j})\mathbf{G}$, то в результате декодирования будет получена информационная последовательность $\mathbf{I} = \mathbf{p}+\Psi_{\ell-j}$. Из линейности кода следует, что $(\mathbf{I}+\Psi_{\ell-j})\mathbf{G} = \mathbf{c}+\Psi_{\ell-j}\mathbf{G}$. По построению кодовое слово $\mathbf{c}$ маскируется ошибкой $\mathbf{e}$, $1 \le wt(\mathbf{e}) \le t_j$, и $\hat{\mathbf{c}} = \mathbf{c}+\mathbf{e}$. Решение j-й подшарады заключается в нахождении ошибки $\mathbf{e}$. Пусть заданы $\hat{\mathbf{c}}$, $\Psi_{\ell-j}$ и некоторая ошибка $\check{\mathbf{e}} \neq \mathbf{e}$. Для $\check{\mathbf{c}} = \hat{\mathbf{c}}+\check{\mathbf{e}}$ в результате декодирования будет получена информационная последовательность $\check{\mathbf{I}} \neq \mathbf{I}$. Решение $\check{\mathbf{e}}$ будет отвергнуто, так как $(\check{\mathbf{I}}+\Psi_{\ell-j})\mathbf{G} \neq \check{\mathbf{c}}+\Psi_{\ell-j}\mathbf{G}$.

Поскольку применяется безызбыточный код, то для j-й подшарады существует q^n кодовых слов. При $1 \le t_j \le n/2$ справедливо неравенство $q^n > \sum_{i=1}^{t_j}(q-1)^i\binom{n}{i}$, и для отыскания решения исчерпывающий перебор ошибок ограниченного веса выгоднее, чем исчерпывающий перебор информационных последовательностей.

Помимо ширины диапазона значение также имеет функция, при помощи которой задается трудоемкость. Для шарад [5] трудоемкость отыскания решения определяется мощностью множества возможных претендентов и равна 2^r для некоторого параметра r. В ряде случаев экспоненциальное изменение трудоемкости не адекватно воздействию и поэтому не оправдано. Необходима плавная регулировка. $RS_q^{\ell}(n, 1)^*$-шарады допускают такую регулировку за счет полиномиальной по n функции изменения трудоемкости для каждой подшарады и общей линейной зависимости. Как было показано (3), трудоемкость отыскания решения для $RS_q^{\ell}(n, 1)^*$-шарады задается аддитивной функцией и допускает гибкую настройку шага изменения трудоемкости. Очевидно, что $\binom{n}{i+1} = \binom{n}{i}\frac{n-i}{i+1}$, $1 \le i \le n$. Тогда при увеличении/уменьшении на единицу веса t ошибки $\mathbf{e}$ трудоемкость отыскания решения возрастает/убывает в $(n-t)/(t+1)$раз. Поскольку $\binom{n}{1} = n$, то для $RS_q^{\ell}(n, 1)^*$-шарады минимальный шаг изменения трудоемкости равен n. Например, можно сконструировать шараду со средней трудоемкостью отыскания решения $\frac{\ell n}{2}$.

Заключение

В статье рассматривается метод шарад как способ противодействия DoS-атаке. Предложены конкретные конструкции шарад на основе кодов, исправляющих ошибки, в том числе и итеративная конструкция, которая гарантирует устойчивость, обладает широким диапазоном трудоемкости и допускает гибкую настройку.

Литература

1. **Naraine R.** Massive DDOS Attack Hit DNS Root Servers. Oct. 23, 2002. http://siliconvalley.internet.com/news/article.php/1486981 (дата обращения: 10.06.2010).
2. **Wagner J.** SCO Hit By Another DDoS Attack. Dec. 10, 2003. http://www.internetnews.com/dev-news/article.php/3287781 (дата обращения: 10.06.2010).

3. **Schuba C. L.** et al. **Analysis of a denial of service attack on TCP** // Proc. IEEE Symp. on Security and Privacy. IEEE Computer Society Press, May 1997. P. 208–223.
4. **Merkle R. C.** Secure Communications Over Insecure Channels // Communications of the ACM. Apr. 1978. Vol. 21. N 4. P. 294–299.
5. **Brainard J., Juels A.** Client Puzzles: A Cryptographic Countermeasure Against Connection Depletion Attacks // Proc. of the 1999 ISOC Network and Distributed System Security Symp. 1999. P. 151–165.
6. **Aura T., Nikander P., Leiwo J.** DoS-resistant authentication with client puzzles // 8th Intern. Workshop on Security Protocols. Springer-Verlag, 2000. P. 170–181.
7. **Dean D., Stubblefield A.** Using client puzzles to protect TLS//SSYM'01 Proc. of the 10th Conf. on USENIX Security Symp./ USENIX Association Berkeley. CA, USA, 2001. P. 1–8.
8. **Adkins D., Lakshminarayanan K., Perrig A., Stoica I.** Taming IP packet flooding attacks // Computer Communication Review. 2004. Vol. 34. N 1. P. 45–50.
9. **Waters B., Juels A., Halderman A., Felten E.** New Client Puzzle Outsourcing Techniques for DoS Resistance // ACM CCS. 2004. P. 246–256.
10. **Rivest R. L., Shamir A., Wagner D.** Time-lock puzzles and timed-release crypto // Technical Report MIT/LCS/TR-684. MIT, 1996. http://people.csail.mit.edu/rivest/RivestShamirWagner-timelock.pdf (дата обращения: 22.10.2010).
11. **Shor P. W.** Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // SIAM J. of Computing. 1997. N 26. P. 1484–1509.
12. **Ekert A., Jozsa R.** Quantum computation and Shor's factoring algorithm // Rev. Mod. Phys. 1996. Vol. 68. N 3. P. 733–753.
13. **Hanneke D.** et al. Realization of a programmable two-qubit quantum processor // Nature Physics. 2009. N 6. P. 13–16.
14. **Vandersypen L. M. K.** et al. Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance // Nature. 2001. N 414. P. 883–887. http://www.nature.com/nature/journal/v414/n6866/abs/414883a.html (дата обращения: 22.10.2010).
15. **Beauregard S.** Circuit for Shor's algorithm using 2n + 3 qubits // Quantum Information and Computation. 2003. N 3. P. 175–185.
16. **Barg A.** Complexity Issues in Coding Theory // Electronic Colloquium on Computational Complexity (ECCC). 1997. Vol. 4. N 46. http://www.eccc.uni-trier.de/report/1997/046/ (дата обращения: 22.10.2010).
17. **Riek J. R.** Observations on the Application of Error-Correcting Codes to Public Key Encryption // Proc. of the IEEE 1990 Intern. Carnahan Conf. on Security Technology, Crime Countermeasures. 1990. P. 15–18.
18. **МакВильямс Ф. Д., Слоэн Н. Дж. Теория кодов, исправляющих ошибки.** — М.: Связь, 1979. — 744 с.
19. **Menezes A., van Oorschot P., Vanstone S.** Handbook of Applied Cryptography. 5th print. — CRC-Press, 2001. — 780 p. http://www.cacr.math.uwaterloo.ca/hac/about/chap9.pdf (дата обращения: 22.10.2010).

УДК 681.5.015

АКТИВНАЯ ПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ СТОХАСТИЧЕСКИХ НЕЛИНЕЙНЫХ НЕПРЕРЫВНО-ДИСКРЕТНЫХ СИСТЕМ НА ОСНОВЕ ЛИНЕАРИЗАЦИИ ВО ВРЕМЕННОЙ ОБЛАСТИ

В. М. Чубич,
канд. техн. наук, доцент
Новосибирский государственный технический университет

Впервые рассмотрены теоретические и прикладные аспекты активной параметрической идентификации стохастических нелинейных непрерывно-дискретных систем. Приведены оригинальные результаты для случая, когда подлежащие оцениванию параметры математических моделей могут входить в уравнения состояния и наблюдения, начальные условия и ковариационные матрицы помех динамики и ошибок измерений. Рассмотрен пример оптимального оценивания параметров одной модельной структуры.

Ключевые слова *— оценивание параметров, метод максимального правдоподобия, планирование оптимальных входных сигналов, информационная матрица Фишера, критерий оптимальности.*

Введение

Проблема идентификации относится к одной из основных проблем теории и практики автоматического управления и является обязательным элементом решения крупномасштабных прикладных задач. Качественное решение данной проблемы способствует эффективному использованию современных математических методов и сложных наукоемких технологий при проектировании различных систем управления подвижными и технологическими объектами, построении прогнозирующих моделей, конструировании следящих и измерительных систем.

По способу проведения эксперимента существующие методы идентификации можно разделить на пассивные и активные. При пассивной идентификации для построения математической модели используются реально действующие в системе сигналы и нормальный режим эксплуатации не нарушается. Методы пассивной идентификации достаточно полно описаны, например, в работе [1]. Активная идентификация, напротив, предполагает нарушение технологического режима и подачу на вход изучаемой системы специальным образом синтезированного сигнала. Его находят в результате решения экстремальной задачи для некоторого предварительно выбранного функционала от информационной (или дисперсионной) матрицы вектора оцениваемых параметров. Трудности, связанные с необходимостью нарушения технологического режима, должны окупаться за счет повышения эффективности и корректности проводимых исследований, что обусловлено самой идеологией активной идентификации, базирующейся на сочетании приемов параметрического оценивания с концепцией планирования эксперимента [2—4].

Более определенно процедура активной идентификации систем с предварительно выбранной модельной структурой предполагает выполнение следующих этапов:

1) вычисление оценок параметров по измерительным данным, соответствующим некоторому пробному сигналу;

2) синтез на основе полученных оценок оптимального по некоторому выбранному критерию сигнала (планирование эксперимента);

3) пересчет оценок неизвестных параметров по измерительным данным, соответствующим синтезированному сигналу.

Целесообразность применения концепции активной идентификации при построении математических моделей стохастических динамических систем показана, например, в работах [5—13]. При этом основное внимание зарубежных уче-

ных в настоящее время обращено на линейные модели в форме передаточных функций [7, 8, 10], нелинейные **FIR-модели [7, 11] и детерминированные** нелинейные модели в пространстве состояний [5, 9]. Стохастические модели в пространстве состояний рассматривались авторами работ [6, 12, 13]. Тем не менее, данная область исследований остается еще недостаточно изученной, а возможности применения в ней методов оптимального планирования экспериментов выявлены далеко не полностью. В настоящей статье приведены результаты исследований автора в рамках указанной проблемы применительно к многомерным стохастическим нелинейным непрерывно-дискретным системам, описываемым моделями в пространстве состояний.

Постановка задачи

Рассмотрим следующую модель управляемой, наблюдаемой, идентифицируемой динамической системы в пространстве состояний:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t)=\mathbf{f}\left[\mathbf{x}(t),\,\mathbf{u}(t),\,t\right]+\mathbf{G}(t)\mathbf{w}(t),\ t\in[t_0,\,t_N]; \quad (1)$$

$$\mathbf{y}(t_{k+1})=\mathbf{h}\left[\mathbf{x}(t_{k+1}),\,t_{k+1}\right]+\mathbf{v}(t_{k+1}),$$
$$k=0,\,1,\,\ldots,\,N-1, \quad (2)$$

где $\mathbf{x}(t)$ — n-вектор состояния; $\mathbf{u}(t)$ — детерминированный r-вектор управления (входа); $\mathbf{w}(t)$ — p-вектор возмущения; $\mathbf{y}(t_{k+1})$ — m-вектор измерения (выхода); $\mathbf{v}(t_{k+1})$ — m-вектор ошибки измерения.

Предположим, что:

• вектор-функции $\mathbf{f}[\mathbf{x}(t),\,\mathbf{u}(t),\,t]$ и $\mathbf{h}[\mathbf{x}(t_{k+1}),\,t_{k+1}]$ непрерывны и дифференцируемы по $\mathbf{x}(t)$, $\mathbf{u}(t)$ и $\mathbf{x}(t_{k+1})$ соответственно; случайные векторы $\mathbf{w}(t)$ и $\mathbf{v}(t_{k+1})$ являются стационарными белыми гауссовыми шумами, для которых

$$\mathrm{E}\left[\mathbf{w}(t)\right]=0,\ \ \mathrm{E}\left[\mathbf{w}(t)\mathbf{w}^T(\tau)\right]=\mathbf{Q}\delta(t-\tau);$$

$$\mathrm{E}\left[\mathbf{v}(t_{k+1})\right]=0,\ \ \mathrm{E}\left[\mathbf{v}(t_{k+1})\mathbf{v}^T(t_{i+1})\right]=\mathbf{R}\delta_{ki};$$

$$\mathrm{E}\left[\mathbf{v}(t_{k+1})\mathbf{w}^T(\tau)\right]=0,\ k,\,i=0,\,1,\,\ldots,\,N-1,$$
$$\tau\in[t_0,\,t_N]$$

(здесь и далее $\mathrm{E}[\cdot]$ — оператор математического ожидания, $\delta(t-\tau)$ — дельта-функция, δ_{ki} — символ Кронекера);

• начальное состояние $\mathbf{x}(t_0)$ имеет нормальное распределение с параметрами

$$\mathrm{E}\left[\mathbf{x}(t_0)\right]=\bar{\mathbf{x}}_0,\ \mathrm{E}\left\{\left[\mathbf{x}(t_0)-\bar{\mathbf{x}}_0\right]\left[\mathbf{x}(t_0)-\bar{\mathbf{x}}_0\right]^T\right\}=\mathbf{P}_0$$

и не коррелирует с $\mathbf{w}(t)$ и $\mathbf{v}(t_{k+1})$ при любых значениях переменной k;

• подлежащие оцениванию параметры $\boldsymbol{\Theta}=(\theta_1,\,\theta_2,\,\ldots,\,\theta_s)$ могут содержаться в вектор-функциях $\mathbf{f}[\mathbf{x}(t),\,\mathbf{u}(t),\,t]$, $\mathbf{h}[\mathbf{x}(t_{k+1}),\,t_{k+1}]$, матрицах $\mathbf{G}(t)$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{P}_0$ и векторе $\bar{\mathbf{x}}_0$ в различных комбинациях.

Необходимо для математической модели (1), (2) с учетом высказанных априорных предположений разработать процедуру активной параметрической идентификации, включающую в себя оценивание параметров и планирование входных сигналов, исследовать эффективность и целесообразность применения указанной процедуры. В такой постановке задача рассматривается и решается впервые.

Линеаризация модели

Считая значение вектора неизвестных параметров $\boldsymbol{\Theta}$ фиксированным, выполним линеаризацию во временной области нелинейной модели (1), (2) относительно номинальной траектории $\{\mathbf{x}_H(t),\,t\in[t_0,\,t_N]\}$, для которой

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_H(t)=\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right],\ t\in[t_0,\ t_N];\\ \mathbf{x}_H(t_0)=\bar{\mathbf{x}}_0.\end{cases} \quad (3)$$

Разложив вектор-функции $\mathbf{f}[\mathbf{x}(t),\ \mathbf{u}(t),\ t]$ и $\mathbf{h}[\mathbf{x}(t_{k+1}),\,t_{k+1}]$ в ряды Тейлора в окрестностях точек $[\mathbf{x}_H(t),\,\mathbf{u}_H(t)]$ и $\mathbf{x}_H(t_{k+1})$ соответственно и отбросив члены второго и более высоких порядков, запишем уравнения линеаризованной модели

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t)=\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]+$$
$$+\frac{\partial\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]}{\partial\mathbf{x}(t)}\left[\mathbf{x}(t)-\mathbf{x}_H(t)\right]+$$
$$+\frac{\partial\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]}{\partial\mathbf{u}(t)}\left[\mathbf{u}(t)-\mathbf{u}_H(t)\right]+\mathbf{G}(t)\mathbf{w}(t); \quad (4)$$

$$\mathbf{y}(t_{k+1})=\mathbf{h}\left[\mathbf{x}_H(t_{k+1}),\,t_{k+1}\right]+$$
$$+\frac{\partial\mathbf{h}\left[\mathbf{x}_H(t_{k+1}),\,t_{k+1}\right]}{\partial\mathbf{x}(t_{k+1})}\times$$
$$\times\left[\mathbf{x}(t_{k+1})-\mathbf{x}_H(t_{k+1})\right]+\mathbf{v}(t_{k+1}), \quad (5)$$

для которой и будем решать поставленную задачу. С учетом обозначений

$$\mathbf{a}(t)=\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]-$$
$$-\frac{\partial\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]}{\partial\mathbf{x}(t)}\mathbf{x}_H(t)+$$
$$+\frac{\partial\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]}{\partial\mathbf{u}(t)}\left[\mathbf{u}(t)-\mathbf{u}_H(t)\right]; \quad (6)$$

$$\mathbf{F}(t)=\frac{\partial\mathbf{f}\left[\mathbf{x}_H(t),\,\mathbf{u}_H(t),\,t\right]}{\partial\mathbf{x}(t)}; \quad (7)$$

$$\mathbf{A}(t_{k+1}) = \mathbf{h}\left[\mathbf{x}_H(t_{k+1}),\ t_{k+1}\right] - \frac{\partial \mathbf{h}\left[\mathbf{x}_H(t_{k+1}),\ t_{k+1}\right]}{\partial \mathbf{x}(t_{k+1})}\mathbf{x}_H(t_{k+1}); \quad (8)$$

$$\mathbf{H}(t_{k+1}) = \frac{\partial \mathbf{h}\left[\mathbf{x}_H(t_{k+1}),\ t_{k+1}\right]}{\partial \mathbf{x}(t_{k+1})} \quad (9)$$

соотношения (4), (5) определяют непрерывно-дискретную модель гауссовой линейной нестационарной системы, описывающейся уравнениями

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{a}(t) + \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{w}(t),\ t \in [t_0,\ t_N]; \quad (10)$$

$$\mathbf{y}(t_{k+1}) = \mathbf{A}(t_{k+1}) + \mathbf{H}(t_{k+1})\mathbf{x}(t_{k+1}) + \mathbf{v}(t_{k+1}),$$

$$k = 0,\ 1,\ \ldots,\ N-1. \quad (11)$$

Заметим, что изложенный способ линеаризации не применим к неоднозначным функциям и нелинейностям, имеющим угловые точки и разрывы. Для линеаризации таких нелинейностей можно воспользоваться методом статистической линеаризации.

Оценивание неизвестных параметров

Оценивание неизвестных параметров математической модели осуществляется по данным наблюдений Ξ в соответствии с критерием идентификации $\chi(\Theta)$. Сбор числовых данных происходит в процессе проведения идентификационных экспериментов, которые выполняются по некоторому плану ξ_ν.

Предположим, что экспериментатор может произвести ν запусков системы, причем сигнал $\mathbf{U}_1$ он подает на вход системы k_1 раз, сигнал $\mathbf{U}_2$ — k_2 раза и т. д., наконец, сигнал $\mathbf{U}_q$ — k_q раз. В этом случае дискретный (точный) нормированный план эксперимента ξ_ν представляет собой совокупность точек $\mathbf{U}_1,\ \mathbf{U}_2,\ \ldots,\ \mathbf{U}_q$, называемых спектром плана, и соответствующих им долей повторных запусков:

$$\xi_\nu = \begin{Bmatrix} \mathbf{U}_1,\ \mathbf{U}_2, \ldots, \mathbf{U}_q \\ \frac{k_1}{\nu}, \frac{k_2}{\nu}, \ldots, \frac{k_q}{\nu} \end{Bmatrix},\ \mathbf{U}_i \in \Omega_{\mathbf{U}},\ \ i = 1,\ 2,\ \ldots,\ q.$$

$\Omega_{\mathbf{U}} \subset \mathbf{R}^{Nr}$ задает ограничения на условия проведения эксперимента. Будем считать, что входной сигнал $\mathbf{u}(t)$ является кусочно-постоянным на фиксированных интервалах времени:

$$\mathbf{u}(t) = \mathbf{u}(t_k),\ t_k \le t \le t_{k+1},\ \ k = 0,\ 1,\ \ldots,\ N-1,$$

т. е. для каждой точки $\mathbf{U}_i$ спектра плана ξ_ν справедливо

$$\mathbf{U}_i^T = \left\{\left[\mathbf{u}^i(t_0)\right]^T, \left[\mathbf{u}^i(t_1)\right]^T, \ldots, \left[\mathbf{u}^i(t_{N-1})\right]^T\right\}.$$

Обозначим через $\mathbf{Y}_{i,j}$ j-ю реализацию выходного сигнала (j=1, 2, ..., k_i), соответствующую i-му входному сигналу $\mathbf{U}_i$ (i=1, 2, ..., q). Тогда в результате проведения по плану ξ_ν идентификационных экспериментов будет сформировано множество

$$\Xi = \left\{\left(\mathbf{U}_i, \mathbf{Y}_{i,j}\right),\ j = 1,\ 2,\ \ldots,\ k_i, i = 1,\ 2,\ \ldots,\ q\right\},$$

$$\sum_{i=1}^{q} k_i = \nu.$$

Уточним структуру $\mathbf{Y}_{i,j}$:

$$\mathbf{Y}_{i,j}^T = \left\{\left[\mathbf{y}^{i,j}(t_1)\right]^T,\ \left[\mathbf{y}^{i,j}(t_2)\right]^T,\ \ldots,\ \left[\mathbf{y}^{i,j}(t_N)\right]^T\right\},$$

$$j = 1,\ 2,\ \ldots,\ k_i,\ i = 1,\ 2,\ \ldots,\ q,$$

и заметим, что в случае пассивной параметрической идентификации, как правило, q=ν=1.

Априорные предположения, высказанные при постановке задачи, и выполненная линеаризация моделей состояния и наблюдения относительно выбранной детерминированной опорной траектории (3) позволяют воспользоваться для оценивания параметров методом максимального правдоподобия. В соответствии с этим методом необходимо найти такие значения параметров $\hat{\Theta}$, для которых

$$\hat{\Theta} = \arg\min_{\Theta \in \Omega_\Theta}\left[-\ln L(\Theta;\Xi)\right] = \arg\min_{\Theta \in \Omega_\Theta}\left[\chi(\Theta)\right], \quad (12)$$

где в соответствии с [1, 14]

$$\chi(\Theta) = \frac{Nm\nu}{2}\ln 2\pi + \frac{1}{2}\sum_{i=1}^{q} k_i \sum_{k=0}^{N-1} \ln\det \mathbf{B}^i(t_{k+1}) + \frac{1}{2} \times$$

$$\times \sum_{i=1}^{q}\sum_{j=1}^{k_i}\sum_{k=0}^{N-1}\left[\boldsymbol{\varepsilon}^{i,j}(t_{k+1})\right]^T\left[\mathbf{B}^i(t_{k+1})\right]^{-1}\left[\boldsymbol{\varepsilon}^{i,j}(t_{k+1})\right], \quad (13)$$

причем

$$\boldsymbol{\varepsilon}^{i,j}(t_{k+1}) = \mathbf{y}^{i,j}(t_{k+1}) - \hat{\mathbf{y}}^{i,j}(t_{k+1}\,|\,t_k),$$

а $\hat{\mathbf{y}}^{i,j}(t_{k+1}\,|\,t_k)$ и $\mathbf{B}^i(t_{k+1})$ определяются по рекуррентным уравнениям непрерывно-дискретного фильтра Калмана (см., например, [15])

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}}^{i,j}(t\,|\,t_k) = \mathbf{F}^i(t)\hat{x}^{i,j}(t\,|\,t_k) + \mathbf{a}^i(t),\quad t_k \le t \le t_{k+1};$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{P}^i(t\,|\,t_k) = \mathbf{F}^i(t)\mathbf{P}^i(t\,|\,t_k) + \mathbf{P}^i(t\,|\,t_k)\left[\mathbf{F}^i(t)\right]^T + \mathbf{G}(t)\mathbf{Q}\mathbf{G}^T(t),$$

$$t_k \le t \le t_{k+1};$$

$$\mathbf{B}^i(t_{k+1}) = \mathbf{H}^i(t_{k+1})\mathbf{P}^i(t_{k+1}\,|\,t_k)\left[\mathbf{H}^i(t_{k+1})\right]^T + \mathbf{R};$$

$$\mathbf{K}^i(t_{k+1}) = \mathbf{P}^i(t_{k+1}\,|\,t_k)\left[\mathbf{H}^i(t_{k+1})\right]^T\left[\mathbf{B}^i(t_{k+1})\right]^{-1};$$

$$\hat{\mathbf{x}}^{i,j}(t_{k+1}\,|\,t_{k+1}) = \hat{\mathbf{x}}^{i,j}(t_{k+1}\,|\,t_k) + \mathbf{K}^i(t_{k+1})\boldsymbol{\varepsilon}^{i,j}(t_{k+1});$$

$$\mathbf{P}^i(t_{k+1}\,|\,t_{k+1}) = \left[\mathbf{I} - \mathbf{K}^i(t_{k+1})\mathbf{H}^i(t_{k+1})\right]\mathbf{P}^i(t_{k+1}\,|\,t_k);$$

$$\hat{\mathbf{y}}^{i,j}(t_{k+1}\,|\,t_k) = \mathbf{A}^i(t_{k+1}) + \mathbf{H}^i(t_{k+1})\hat{\mathbf{x}}^{i,j}(t_{k+1}\,|\,t_k)$$

для $k=0, 1, ..., N-1$, $j = 1, 2, ..., k_i$, $i = 1, 2, ..., q$ начальными условиями $\hat{\mathbf{x}}(t_0\,|\,t_0) = \bar{\mathbf{x}}_0$, $\mathbf{P}(t_0, t_0) = \mathbf{P}_0$.

Для нахождения условного минимума $\chi(\boldsymbol{\Theta})$ воспользуемся методом проекции градиента [16], учитывая, что

$$\frac{\partial\chi(\boldsymbol{\Theta})}{\partial\theta_l} = \sum_{i=1}^{q}\sum_{j=1}^{k_i}\sum_{k=0}^{N-1}\left\{\left[\frac{\partial\boldsymbol{\varepsilon}^{i,j}(t_{k+1})}{\partial\theta_l}\right]^T\left[\mathbf{B}^i(t_{k+1})\right]^{-1}\times\right.$$
$$\times\left[\boldsymbol{\varepsilon}^{i,j}(t_{k+1})\right] - \frac{1}{2}\left[\boldsymbol{\varepsilon}^{i,j}(t_{k+1})\right]^T\left[\mathbf{B}^i(t_{k+1})\right]^{-1}\times$$
$$\left.\times\frac{\partial\mathbf{B}^i(t_{k+1})}{\partial\theta_l}\left[\mathbf{B}^i(t_{k+1})\right]^{-1}\boldsymbol{\varepsilon}^{i,j}(t_{k+1})\right\} +$$
$$+\frac{1}{2}\sum_{i=1}^{q}k_i\sum_{k=0}^{N-1}\mathrm{Sp}\left\{\left[\mathbf{B}^i(t_{k+1})\right]^{-1}\frac{\partial\mathbf{B}^i(t_{k+1})}{\partial\theta_l}\right\},$$
$$l = 1, 2, \ldots, s.$$

Здесь частные производные $\dfrac{\partial\boldsymbol{\varepsilon}^{i,j}(t_{k+1})}{\partial\theta_l}$ и $\dfrac{\partial\mathbf{B}^i(t_{k+1})}{\partial\theta_l}$ по аналогии с [12] вычисляются по рекуррентным аналитическим формулам, вытекающим из уравнений непрерывно-дискретного фильтра Калмана.

Планирование входных сигналов

Предварим рассмотрение алгоритмов синтеза оптимальных входных сигналов изложением некоторых основополагающих понятий и результатов теории планирования эксперимента для нашего случая.

Под непрерывным нормированным планом $\boldsymbol{\xi}$ условимся понимать совокупность величин

$$\boldsymbol{\xi} = \begin{Bmatrix}\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_q\\ p_1, p_2, \ldots, p_q\end{Bmatrix}, \quad p_i \ge 0, \ \sum_{i=1}^{q} p_i = 1,$$

$$\mathbf{U}_i \in \boldsymbol{\Omega}_{\mathbf{U}}, \quad i = 1, 2, \ldots, q. \qquad (14)$$

Здесь точки спектра $\mathbf{U}_i$ имеют такую же структуру, как и в случае дискретного плана $\boldsymbol{\xi}_N$, но веса p_i могут принимать любые значения в диапазоне от 0 до 1, в том числе и иррациональные. Множество планирования $\boldsymbol{\Omega}_{\mathbf{U}}$ определяется ограничениями на условия проведения эксперимента.

Для плана (14) нормированная информационная матрица $\mathbf{M}(\boldsymbol{\xi})$ определяется соотношением

$$\mathbf{M}(\boldsymbol{\xi}) = \sum_{i=1}^{q} p_i\mathbf{M}(\mathbf{U}_i;\,\boldsymbol{\Theta}), \qquad (15)$$

в котором информационные матрицы Фишера одноточечных планов (индекс i для простоты записи соотношения опустим)

$$\mathbf{M}(\mathbf{U};\,\boldsymbol{\Theta}) = -\mathrm{E}_{\mathbf{Y}}\left[\frac{\partial^2\ln L\left(\boldsymbol{\Theta};\,\mathbf{Y}_1^N\right)}{\partial\boldsymbol{\Theta}\partial\boldsymbol{\Theta}^T}\right]$$

зависят от неизвестных параметров $\boldsymbol{\Theta}$, что позволяет в дальнейшем говорить только о локально-оптимальном планировании. Получено [17] весьма сложное в математическом отношении и громоздкое для представления в рамках данной статьи выражение для информационных матриц Фишера $\mathbf{M}(\mathbf{U};\,\boldsymbol{\Theta})$, соответствующее модели (10), (11).

Качество оценивания параметров моделей можно повысить за счет построения плана эксперимента, оптимизирующего некоторый выпуклый функционал X от информационной матрицы $\mathbf{M}(\boldsymbol{\xi})$ путем решения экстремальной задачи

$$\boldsymbol{\xi}^* = \arg\min_{\boldsymbol{\xi}\in\boldsymbol{\Omega}_{\boldsymbol{\xi}}} X[\mathbf{M}(\boldsymbol{\xi})]. \qquad (16)$$

Решая задачу планирования эксперимента, мы определенным образом воздействуем на нижнюю границу неравенства Рао—Крамера [1]: например, для **D-оптимального плана минимизируем** объем, для A-оптимального — сумму квадратов длин осей эллипсоида рассеяния оценок параметров.

При решении экстремальной задачи (16) возможны два подхода. Первый из них (прямой) предполагает поиск минимума функционала $X[\mathbf{M}(\boldsymbol{\xi})]$ непосредственно с привлечением методов нелинейного программирования. Возможные варианты прямой процедуры синтеза оптимальных входных сигналов представлены в работах [3, 6, 12, 13]. Другой подход (его называют двойственным) основан на теореме эквивалентности [18], обобщенная формулировка которой выглядит следующим образом.

Утверждения:

1) план $\boldsymbol{\xi}^*$ минимизирует $X[\mathbf{M}(\boldsymbol{\xi})]$;

2) план $\boldsymbol{\xi}^*$ минимизирует $\max\limits_{\mathbf{U}\in\boldsymbol{\Omega}_{\mathbf{U}}}\mu(\mathbf{U},\boldsymbol{\xi})$;

3) $\max\limits_{\mathbf{U}\in\boldsymbol{\Omega}_{\mathbf{U}}}\mu\left(\mathbf{U},\,\boldsymbol{\xi}^*\right) = \eta$

эквивалентны между собой. Информационные матрицы планов, удовлетворяющих условиям 1—3, совпадают. Любая линейная комбинация планов, удовлетворяющих 1—3, также удовлет-

■ ***Таблица 1.*** *Соответствие значений параметров теоремы эквивалентности критериям оптимальности*

Критерий оптимальности	Параметры теоремы эквивалентности		
	$X[\mathbf{M}(\boldsymbol{\xi})]$	$\mu(\mathbf{U}, \boldsymbol{\xi})$	η
D	$-\ln\det\mathbf{M}(\boldsymbol{\xi})$	$\mathrm{Sp}\left[\mathbf{M}^{-1}(\boldsymbol{\xi})\mathbf{M}(\mathbf{U})\right]$	s
A	$\mathrm{Sp}\left[\mathbf{M}^{-1}(\boldsymbol{\xi})\right]$	$\mathrm{Sp}\left[\mathbf{M}^{-2}(\boldsymbol{\xi})\mathbf{M}(\mathbf{U})\right]$	$\mathrm{Sp}\left[\mathbf{M}^{-1}(\boldsymbol{\xi})\right]$

воряет 1—3. Выражения для $\mu(\mathbf{U}, \boldsymbol{\xi})$, η приведены в табл. 1.

Приведем двойственную градиентную процедуру построения непрерывных оптимальных планов [2, 12, 18].

Шаг 1. Зададим начальный невырожденный план $\boldsymbol{\xi}_0$ и по формуле (15) вычислим нормированную матрицу $\mathbf{M}(\boldsymbol{\xi}_0)$ плана. Положим $l=0$.

Шаг 2. Найдем локальный максимум

$$\mathbf{U}^l = \arg\max_{\mathbf{U}\in\boldsymbol{\Omega}_\mathbf{U}} \mu(\mathbf{U}, \boldsymbol{\xi}_l)$$

методом проекции градиента. Если окажется, что $\left|\mu\left(\mathbf{U}^l, \boldsymbol{\xi}_l\right)-\eta\right|\le\delta$, закончим процесс. Если $\mu\left(\mathbf{U}^l, \boldsymbol{\xi}_l\right)>\eta$, перейдем к шагу 3. В противном случае будем искать новый локальный максимум.

Шаг 3. Вычислим τ_l по формуле

$$\tau_l = \arg\min_{0\le\tau\le1} X\left[\mathbf{M}\left(\boldsymbol{\xi}_{l+1}^{\tau}\right)\right],$$

$$\boldsymbol{\xi}_{l+1}^{\tau} = (1-\tau)\boldsymbol{\xi}_l + \tau\boldsymbol{\xi}\left(\mathbf{U}^l\right),$$

где $\boldsymbol{\xi}(\mathbf{U}^l)$ — одноточечный план, размещенный в точке $\mathbf{U}^l$.

Шаг 4. Составим план $\boldsymbol{\xi}_{l+1} = (1-\tau_l)\boldsymbol{\xi}_l + \tau_l\boldsymbol{\xi}\left(\mathbf{U}^l\right)$, произведем его «очистку» в соответствии с рекомендациями [2], положим $l=l+1$ и перейдем на шаг 2[1].

Приведенный алгоритм построения оптимальных сигналов требует вычисления градиента

$$\nabla_\mathbf{U}\mu(\mathbf{U},\boldsymbol{\xi}) = \left\|\frac{\partial\mu(\mathbf{U},\boldsymbol{\xi})}{\partial u_\alpha\left(t_\beta\right)}\right\|,$$

$$\beta = 0, 1, \ldots, N-1, \ \alpha = 1, 2, \ldots, r.$$

Для критерия D-оптимальности получаем

$$\frac{\partial\mu(\mathbf{U},\boldsymbol{\xi})}{\partial u_\alpha\left(t_\beta\right)} = \frac{\partial\mathrm{Sp}\left[\mathbf{M}^{-1}(\boldsymbol{\xi})\mathbf{M}(\mathbf{U})\right]}{\partial u_\alpha\left(t_\beta\right)} = \mathrm{Sp}\left[\mathbf{M}^{-1}(\boldsymbol{\xi})\frac{\partial\mathbf{M}(\mathbf{U})}{\partial u_\alpha\left(t_\beta\right)}\right].$$

[1] Соответствие значений параметров $X[\mathbf{M}(\boldsymbol{\xi})]$, $\mu(\mathbf{U}, \boldsymbol{\xi})$, η двойственной процедуры критериям A- и D-оптимальности такое же, как и в табл. 1.

В случае критерия A-оптимальности

$$\frac{\partial\mu(\mathbf{U},\boldsymbol{\xi})}{\partial u_\alpha\left(t_\beta\right)} = \frac{\partial\mathrm{Sp}\left[\mathbf{M}^{-2}(\boldsymbol{\xi})\mathbf{M}(\mathbf{U})\right]}{\partial u_\alpha\left(t_\beta\right)} = \mathrm{Sp}\left[\mathbf{M}^{-2}(\xi)\frac{\partial\mathbf{M}(\mathbf{U})}{\partial u_\alpha\left(t_\beta\right)}\right].$$

Основу рассмотренной процедуры синтеза входных сигналов составляют достаточно сложные объемные алгоритмы вычисления информационной матрицы одноточечного плана $\mathbf{M}(\mathbf{U}; \boldsymbol{\Theta})$ и ее производных $\frac{\partial\mathbf{M}(\mathbf{U};\boldsymbol{\Theta})}{\partial u_\alpha\left(t_\beta\right)}$, обстоятельно изложенные автором [17, 19].

Практическое применение в процедуре активной параметрической идентификации построенного непрерывного оптимального плана

$$\boldsymbol{\xi}^* = \begin{Bmatrix}\mathbf{U}_1^*, \mathbf{U}_2^*, \ldots, \mathbf{U}_q^* \\ p_1^*, p_2^*, \ldots, p_q^*\end{Bmatrix}, \ \sum_{i=1}^{q} p_i^* = 1, \ p_i^* \ge 0, \ \mathbf{U}_i^* \in \boldsymbol{\Omega}_\mathbf{U},$$

$$i = 1, 2, \ldots, q$$

затруднено тем обстоятельством, что веса p_i^* представляют собой, вообще говоря, произвольные вещественные числа, заключенные в интервале от 0 до 1. Несложно заметить, что в случае заданного числа ν возможных запусков системы величины $k_i^* = \nu p_i^*$ могут оказаться нецелыми числами. Проведение эксперимента требует округления величин k_i^* до целых чисел. Очевидно, что полученный в результате такого округления план будет отличаться от оптимального непрерывного плана, причем приближение тем лучше, чем больше число возможных запусков. Возможный алгоритм «округления» непрерывного плана до дискретного (точного) изложен в работе [4].

Разработанный в рамках системы MATLAB программный комплекс включает в себя модули, отвечающие за вычисление информационной матрицы и ее производных по компонентам входного сигнала, нахождение оценок неизвестных параметров методом максимального правдоподобия, синтез A- и **D-оптимальных входных сигналов** с использованием прямой и двойственной градиентных процедур.

Пример активной параметрической идентификации

Рассмотрим следующую модель стохастической нелинейной непрерывно-дискретной системы:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -\frac{\theta_2}{\theta_1}x(t) + \frac{0{,}01}{\theta_1}(u(t)-x(t))\times$$

$$\times\exp\left[0{,}25(u(t)-x(t))\right] + \frac{0{,}1}{\theta_1}w(t), \ t\in[t_0, t_N]; \qquad (17)$$

$$y(t_{k+1}) = x(t_{k+1}) + v(t_{k+1}),\ \ k = 0,\ 1,\ \ldots,\ N-1, \quad (18)$$

где θ_1, θ_2 — неизвестные параметры системы, причем $2 \le \theta_1 \le 10$; $0{,}05 \le \theta_2 \le 2$.

Будем считать, что выполнены все априорные предположения, высказанные при постановке задачи, причем

$$\mathrm{E}[w(t)w(\tau)] = 0{,}8\delta(t-\tau) = Q\delta(t-\tau);$$

$$\mathrm{E}[v(t_{k+1})v(t_{i+1})] = 0{,}4\delta_{ki} = R\delta_{ki};$$

$$x(t_0) \in N(0;\ 0{,}01).$$

Выполнив линеаризацию модели (17), (18) во временной области относительно номинальной траектории:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}x_H(t) = -\dfrac{\theta_2}{\theta_1}x_H(t) + \dfrac{0{,}01}{\theta_1}(u_H(t) - x_H(t)) \times \\ \times \exp[0{,}25(u_H(t) - x_H(t))],\ t \in [t_0,\ t_N]; \\ x_H(t_0) = 0, \end{cases} \quad (19)$$

получим линеаризованную модель вида (10), (11), в которой

$$a(t) = \frac{0{,}01}{\theta_1}\exp[0{,}25(u_H(t) - x_H(t))] \times$$
$$\times \{[1 + 0{,}25(u_H(t) - x_H(t))] \times$$
$$\times u(t) - 0{,}25(u_H(t) - x_H(t))^2\};$$

$$F(t) = -\frac{\theta_2}{\theta_1} - \frac{0{,}01}{\theta_1}\exp[0{,}25(u_H(t) - x_H(t))] \times$$
$$\times [1 + 0{,}25(u_H(t) - x_H(t))];$$

$$G(t) = \frac{0{,}1}{\theta_1};\ \ A(t_{k+1}) = 0;\ \ H(t_{k+1}) = 1.$$

Необходимо оценить параметры θ_1, θ_2, входящие в $a(t)$, $F(t)$ и $G(t)$.

Считая, что для номинальной траектории (19) $u_H(t) = u(t)$, $t \in [t_0,\ t_N]$, обеспечим наилучшее приближение построенной линеаризованной модели к своему нелинейному аналогу. Выберем область планирования $\Omega_{\mathbf{U}} = \{\mathbf{U} \in R^N \mid \ 10 \le u(t_k) \le 15,\ k = 0,\ 1,\ \ldots,\ N-1\}$ и критерий D-оптимальности. Для того чтобы ослабить зависимость результатов оценивания от выборочных данных, произведем шесть независимых запусков системы и усредним полученные оценки неизвестных параметров. Реализации выходных сигналов получим компьютерным моделированием при истинных значениях параметров $\theta_1^* = 4$, $\theta_2^* = 0{,}5$ и $t_0 = 0$, $t_N = 30$, $N = 31$.

О качестве идентификации в пространстве параметров и в пространстве откликов будем судить, соответственно, по значениям коэффициентов k_θ и k_Y, вычисляющихся по следующим формулам:

$$k_{\boldsymbol{\theta}} = \frac{\left\|\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}_{\mathrm{cp}}\right\|}{\left\|\boldsymbol{\theta}^* - \hat{\boldsymbol{\theta}}^*_{\mathrm{cp}}\right\|} = \sqrt{\frac{\left(\theta_1^* - \hat{\theta}_{1\mathrm{cp}}\right)^2 + \left(\theta_2^* - \hat{\theta}_{2\mathrm{cp}}\right)^2}{\left(\theta_1^* - \hat{\theta}^*_{1\mathrm{cp}}\right)^2 + \left(\theta_2^* - \hat{\theta}^*_{2\mathrm{cp}}\right)^2}};$$

$$k_{\mathbf{Y}} = \frac{\left\|\mathbf{Y}_{\mathrm{cp}} - \hat{\mathbf{Y}}_{\mathrm{cp}}\right\|}{\left\|\mathbf{Y}_{\mathrm{cp}} - \hat{\mathbf{Y}}^*_{\mathrm{cp}}\right\|} = \sqrt{\frac{\sum_{k=0}^{N-1}\left(y_{\mathrm{cp}}(t_{k+1}) - \hat{y}_{\mathrm{cp}}(t_{k+1} \mid t_{k+1})\right)^2}{\sum_{k=0}^{N-1}\left(y_{\mathrm{cp}}(t_{k+1}) - \hat{y}^*_{\mathrm{cp}}(t_{k+1} \mid t_{k+1})\right)^2}},$$

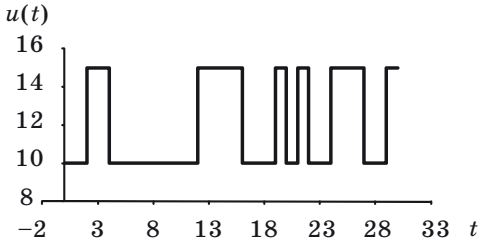
где $\boldsymbol{\theta}^*$ — вектор истинных значений параметров; $\hat{\boldsymbol{\theta}}_{\mathrm{cp}}$ — вектор усредненных оценок неизвестных значений параметров по исходному входному сигналу; $\hat{\boldsymbol{\theta}}^*_{\mathrm{cp}}$ — вектор усредненных оценок неизвестных значений параметров по синтезированному входному сигналу; $\mathbf{Y}_{\mathrm{cp}} = \{y_{\mathrm{cp}}(t_{k+1}),\ k = 0,\ 1,\ \ldots,\ N-1\}$, $\hat{\mathbf{Y}}_{\mathrm{cp}} = \{\hat{y}_{\mathrm{cp}}(t_{k+1} \mid t_{k+1}),\ k = 0,\ 1,\ \ldots,\ N-1\}$, $\hat{\mathbf{Y}}^*_{\mathrm{cp}} = \{\hat{y}^*_{\mathrm{cp}}(t_{k+1} \mid t_{k+1}),\ k = 0,\ 1,\ \ldots,\ N-1\}$ — усредненные по всем запускам последовательности измерений для вектора $\boldsymbol{\theta}$, равного $\boldsymbol{\theta}^*$, $\hat{\boldsymbol{\theta}}_{\mathrm{cp}}$, $\hat{\boldsymbol{\theta}}^*_{\mathrm{cp}}$ соответственно, при некотором выбранном допустимом входном сигнале $\mathbf{u}(t) \in \boldsymbol{\Omega}_{\mathbf{U}}$; $\hat{y}(t_{k+1} \mid t_{k+1})$ находится при помощи равенства

$$\hat{y}(t_{k+1} \mid t_{k+1}) = A(t_{k+1}) + H(t_{k+1})\hat{x}(t_{k+1} \mid t_{k+1}).$$

Результаты выполнения процедуры активной параметрической идентификации представлены в табл. 2 (оптимальный план получился одноточечным).

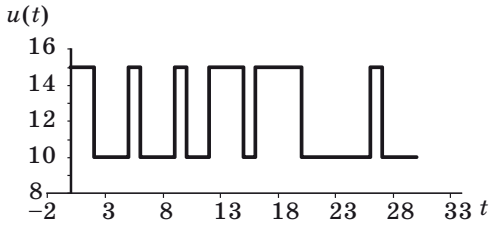
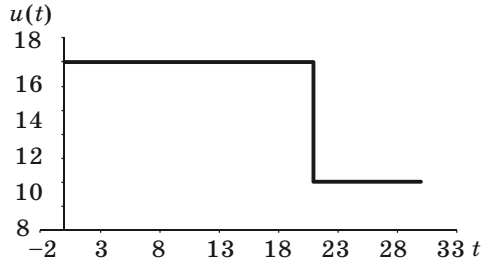
Данные табл. 2 показывают, что коэффициент $k_{\boldsymbol{\theta}} \approx 5{,}27$. В пространстве откликов при псевдослучайном входном сигнале $\mathbf{u}(t)$, представленном на рис. 1, $k_{\mathbf{Y}} \approx 1{,}13$. При решении реальных задач истинные значения параметров неизвестны и, таким образом, сравнение качества оценок в пространстве параметров невозможно. Именно поэтому наиболее показательным является сравнение качества оценивания в пространстве откликов (рис. 2, *а*, *б*).

Значения полученных коэффициентов k_θ и k_Y позволяют сделать вывод об эффективности и це-

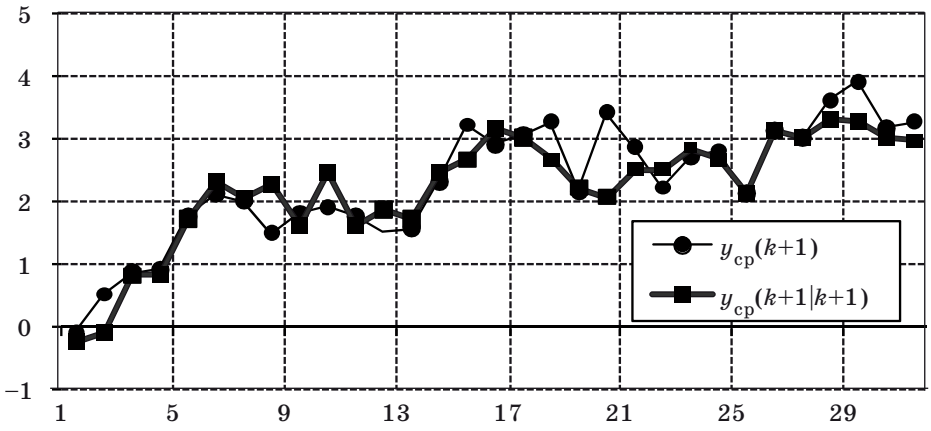


■ ***Рис. 1.*** *Тестовый сигнал для анализа качества прогнозирования на основе результатов из табл. 2*

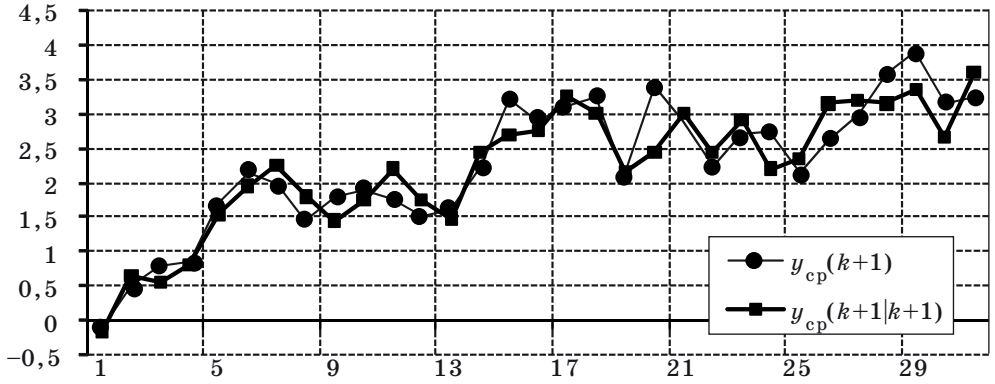
■ ***Таблица 2.*** *Результаты выполнения процедуры активной идентификации*

Исходный и синтезированный входные сигналы	Номер запуска системы	Значения оценок параметра	
		$\hat{\theta}_1$	$\hat{\theta}_2$
	1	3,587	0,434
	2	6,099	0,565
	3	5,735	0,486
	4	3,158	0,504
	5	3,354	0,484
	6	4,002	0,413
	Средние значения по запускам	4,322	0,481
	1	3,156	0,522
	2	4,724	0,587
	3	4,917	0,537
	4	3,320	0,460
	5	3,865	0,463
	6	4,375	0,500
	Средние значения по запускам	4,060	0,512

а)



б)



■ ***Рис. 2.*** *Графическое представление* $\mathbf{Y}_{ср}$ *и* $\hat{\mathbf{Y}}^{*}_{ср}$ *при* $u(t)$*, изображенном на рис. 1:* *а* — $y_{ср}(k+1|k+1)$ *соответствует* $\hat{y}_{ср}(t_{k+1}|t_{k+1})$; *б* — $y_{ср}(k+1|k+1)$ *соответствует* $\hat{y}^{*}_{ср}(t_{k+1}|t_{k+1})$.

лесообразности применения разработанной процедуры активной параметрической идентификации при построении моделей стохастических нелинейных непрерывно-дискретных систем.

Заключение

Дано систематическое изложение наиболее существенных для практики вопросов теории и техники активной параметрической идентификации стохастических нелинейных непрерывно-дискретных систем. Рассмотрена и решена задача оптимального оценивания на основе линеаризации во временной области для случая вхождения неизвестных параметров в уравнения состояния и наблюдения, начальные условия и ковариационные матрицы помех динамики и ошибок измерений. Разработаны оригинальные градиентные алгоритмы активной идентификации, позволяющие решать задачи оптимального оценивания параметров математических моделей методом максимального правдоподобия с привлечением прямой и двойственной процедур синтеза A- и D-оптимальных входных сигналов. Показана эффективность и целесообразность применения разработанной процедуры активной параметрической идентификации при построении моделей стохастических нелинейных непрерывно-дискретных систем.

Работа выполнена при поддержке Федерального агентства по образованию в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009—2013 гг.

Литература

1. **Льюнг Л.** Идентификация систем: Теория для пользователя. — М.: Наука, 1991. — 432 с.
2. **Федоров В. В.** Теория оптимального эксперимента (планирование регрессионных экспериментов). — М.: Наука, 1971. — 312 с.
3. **Денисов В. И.** Математическое обеспечение системы ЭВМ — экспериментатор. — М.: Наука, 1977. — 251 с.
4. **Ермаков С. М., Жиглявский А. А.** Математическая теория оптимального эксперимента. — М.: Наука, 1987. — 320 с.
5. **Zhao J., Kanellakopoulos I.** Active identification for discrete-time nonlinear control. Part I: Output-feedback systems // IEEE Trans. Automat. Control. 2002. Vol. 47. N 2. P. 210—240.
6. **Денисов В. И., Чубич В. М., Черникова О. С.** Активная параметрическая идентификация стохастических линейных дискретных систем во временной области// Сиб. журн. индустр. матем. 2003. Т. 6. № 3(15). С. 70—87.
7. **Hjalmarsson H.** From experiment design to closed-loop control // Automatica. 2005. Vol. 41. P. 393—438.
8. **Gevers M., Mišković L., Bonvin D., Karimi A.** Identification of multi-input systems: variance analysis and input design issues // Automatica. 2006. Vol. 42. P. 559—572.
9. **Jauberthie C., Denis-Vidal L., Coton P., Joly-Blanchard G.** An optimal input design procedure// Automatica. 2006. Vol. 42. P. 881—884.
10. **Rojas C. R., Welsh J. S., Goodwin G. C., Feuer A.** Robust optimal experiment design for system identification // Automatica. 2007. Vol. 43. P. 993—1008.
11. **Hjalmarsson H., Martensson J., Ninness B.** Optimal input design for identification of nonlinear systems: learning from the linear case // American control conf. (ACC), 9—13 July 2007. P. 1572—1576.
12. **Денисов В. И.** и др. Активная параметрическая идентификация стохастических линейных систем. — Новосибирск: Изд-во НГТУ, 2009. — 192 с.
13. **Chubich V. M.** Application of methods of experiment design theory in problem of stochastic nonlinear discrete systems identification // Proc. of the IASTED international conferences on automation, control, and information technology (ACITCDA 2010), Novosibirsk, Russia, 15—18 June 2010. P. 272—279.
14. **Åström K. J.** Maximum likelihood and prediction errors methods // Automatica. 1980. Vol. 16. P. 551—574.
15. **Огарков М. А.** Методы статистического оценивания параметров случайных процессов. — М.: Энергоатомиздат, 1980. — 208 с.
16. **Базара М., Шетти К.** Нелинейное программирование. — М.: Мир, 1982. — 583 с.
17. **Чубич В. М.** Особенности вычисления информационной матрицы Фишера в задаче активной параметрической идентификации стохастических нелинейных непрерывно-дискретных систем // Науч. вест. НГТУ. 2009. № 1(34). С. 41—54.
18. **Mehra R. K.** Optimal input signals for parameter estimation in dynamic systems — survey and new results // IEEE Trans. Automat. Control. 1974. Vol. 19. N 6. P. 753—768.
19. **Чубич В. М., Филиппова Е. В.** Вычисление производных информационной матрицы Фишера по компонентам входного сигнала в задаче активной параметрической идентификации стохастических нелинейных непрерывно-дискретных систем // Науч. вест. НГТУ. 2010. № 2(39). С. 53—63.

УДК 004.728.3.057.4

АНАЛИЗ ВЛИЯНИЯ ИЗМЕНЕНИЯ ХАРАКТЕРИСТИК ПОТОКА НА ЭНЕРГОЗАТРАТЫ МОБИЛЬНОЙ СТАНЦИИ

А. В. Анисимов,
аспирант
А. М. Тюрликов,
канд. техн. наук, доцент
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассматривается режим ожидания стандарта IEEE 802.16m, приводится анализ средней задержки и показателей энергоэффективности мобильной станции при приеме потока данных с переменной интенсивностью при использовании данного режима.

Ключевые слова *— сбережение энергии, качество обслуживания, режим ожидания.*

Введение

В настоящее время при проектировании и развертывании систем передачи информации все большее распространение получают беспроводные технологии широкополосного доступа. Среди них лидирующие позиции прочно удерживают IEEE 802.16 и LTE (*Long Term Evolution*). Обе технологии обладают широкими возможностями по поддержке работы мобильных пользователей. Поскольку мобильные станции (МС) пользователей имеют ограниченный запас аккумуляторной батареи, остро встает задача выработки новых и улучшения существующих механизмов сбережения энергии, потребляемой МС.

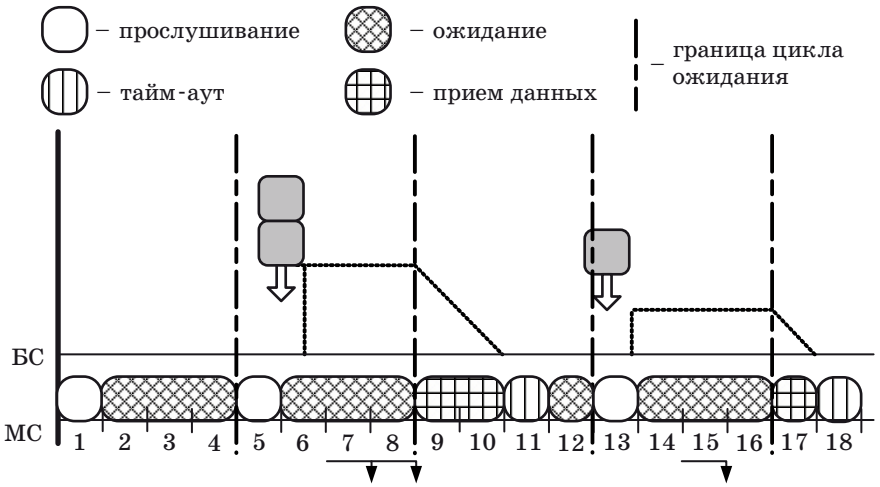
Вместе с тем современные стандарты широкополосных беспроводных технологий ограничиваются общим описанием механизмов сбережения энергии и не предоставляют информации о методах выбора параметров для их оптимальной работы. Существенным также является тот факт, что ненадлежащее использование механизмов сбережения энергии не только повышает энергозатраты МС, но и может привести к снижению показателей качества обслуживания пользователей. Как следствие, важной прикладной задачей становится разработка способов выбора оптимальных параметров для эффективного функционирования механизмов сбережения энергии. Решению данной задачи и посвящена настоящая работа.

Поскольку технология беспроводного доступа IEEE 802.16 к настоящему моменту получила наиболее широкое распространение, дальнейшее изложение построено с учетом ее особенностей согласно новейшему стандарту IEEE 802.16m [1]. Сами особенности стандарта IEEE 802.16m были подробно описаны в книге [2]. В данной работе рассматриваются только аспекты, связанные с функционированием механизма сбережения энергии (или режима ожидания, *sleep mode*) при нисходящей передаче информации.

Режим ожидания стандарта IEEE 802.16m

Согласно стандарту IEEE 802.16m, время функционирования системы разбивается на одинаковые временные отрезки, которые называются кадрами. Далее длительность всех рассматриваемых временных интервалов мы будем измерять в кадрах.

Основная идея режима ожидания состоит в том, что все время функционирования МС разбивается на интервалы времени, которые называются циклами ожидания. Длительность цикла ожидания обозначается C. Каждый цикл ожидания состоит из активного интервала и интервала ожидания. Во время активного интервала радиотракт на МС включен и МС слушает радиоканал, а во время интервала ожидания МС отключает свой радиоприемник и канал не слушает, таким образом снижая свои энергозатраты. Каждый цикл ожидания начинается с активного интервала. На рис. 1 представлен пример функционирования системы в режиме ожи-



■ ***Рис. 1.*** *Режим ожидания*

дания. В начале каждого цикла ожидания в течение некоторого времени базовая станция *БС* сообщает МС, есть ли у нее данные для передачи. Такой интервал времени называется интервалом прослушивания, его длительность обозначается *L*. Если данных нет, МС переключается в состояние ожидания, выключая свой приемник. В случае, когда у БС есть данные для передачи, начинается их передача к МС. При этом длительность активного интервала увеличивается, а интервала ожидания уменьшается. Период времени, в течение которого МС осуществляет прием данных от БС, будем называть интервалом приема. Интервал приема может длиться до конца цикла ожидания.

После каждого кадра, в котором была передача данных, МС продолжает слушать канал в течение некоторого времени. Такой интервал времени будем называть тайм-аутом и обозначать *T*. Наличие тайм-аута необходимо для того, чтобы узнать, есть ли на БС еще данные для передачи. Если за время тайм-аута новой передачи данных не было, МС переключается в состояние ожидания до конца текущего цикла.

Соответственно, режим ожидания полностью задается тремя параметрами: *C, L, T*.

Обзор работ по режиму ожидания

В настоящее время множество научных работ посвящено анализу режима энергосбережения в стандарте IEEE 802.16e. Наиболее полный анализ, проведенный в работах [3, 4], основан на изучении системы M/GI/1/K с перерывами. Авторы приводят также алгоритм оптимизации для случая, когда имеется ограничение на вероятность потери сообщения с данными. В статье [5] детальный анализ режима ожидания построен на изучении системы M/G/1 с перерывами, длительность которых может иметь различные распределения. Помимо анализа приводится набор оптимизационных алгоритмов в зависимости от того, какие параметры системы известны. Однако во всех вышеперечисленных работах в качестве входного потока используется поток Пуассона. Отдельный интерес представляет рассмотрение входных потоков, отличных от пуассоновского. В частности, в работе [6] изучается влияние вида входного потока на энергопотребление МС в режиме ожидания и сравниваются результаты для случаев с пуассоновским входным потоком и входным потоком Эрланга. Анализ задержки и энергопотребления в режиме ожидания при использовании модели входного потока DBMAP (*Discrete-time Batch Markovian Arrival Process*) представлен в работе [7].

Все приведенные работы посвящены изучению режима ожидания стандарта IEEE 802.16e. Данный режим сильно отличается от режима ожидания стандарта IEEE 802.16m. В работе [8] можно найти анализ энергопотребления МС при режиме ожидания стандарта IEEE 802.16m и передаче пуассоновского потока.

В настоящей статье представлен анализ средней задержки и энергозатрат МС во время функционирования в режиме ожидания стандарта IEEE 802.16m при передаче пуассоновского потока и частного случая DBMAP-потока. Кроме этого, ставится и решается оптимизационная задача по обеспечению требуемых параметров качества обслуживания при минимизации энергозатрат МС, а также рассматривается влияние различных моделей потоков данных и изменения их характеристик на среднюю задержку и энергозатраты.

Оптимизационная задача

Технология IEEE 802.16 обеспечивает качество обслуживания (QoS) для различных типов пользовательских потоков. К параметрам качества обслуживания относятся задержка, вариация задержки (*jitter*), минимальная зарезервированная и максимальная скорости потока данных. Очевидно, что использование режима ожидания негативно влияет (увеличивает значения) на задержку и вариацию задержки. Соответственно выбор параметров режима ожидания необходимо осуществлять так, чтобы обеспечить требуемые значения параметров качества обслуживания. В данной работе учитывается только один параметр — средняя задержка. И задача минимизации энергозатрат решается при ограничении на среднюю задержку передачи данных.

Минимизировать

$$f_E(C, L, T)$$

при ограничении

$$f_D(C, L, T) \le D_{\max},$$

где $D_{\max}$ — максимально допустимая средняя задержка; C — длительность цикла ожидания; L — длительность периода прослушивания; T — тайм-аут.

Модель системы передачи информации

Для анализа режима ожидания была разработана модель системы передачи информации, которая описывается следующим набором допущений.

Допущение 1. Рассматривается функционирование системы связи, состоящей из одной БС и одной МС.

Допущение 2. Рассматривается только нисходящая передача данных (от БС к МС).

Допущение 3. В течение кадра может быть передано не более одного сообщения с данными.

Допущение 4. Сообщения передаются на МС в порядке их поступления на БС, без перерывов в расписании.

Допущение 5. Если сообщение с данными поступило в буфер БС в течение кадра с номером k, то оно может быть передано МС не ранее следующего кадра, т. е. кадра с номером $k + 1$.

Допущение 6. Входной поток сообщений представляет собой поток Пуассона.

Допущение 7. Мощность, потребляемая МС в состоянии ожидания, равна P_S, а в состоянии прослушивания — P_A.

Кратко поясним приведенные допущения. Допущение 1 вызвано тем, что влияние расписания передач одной МС на расписания передач других МС определяется набором правил и алгоритмов, которые используются планировщиком на БС. Однако алгоритм работы планировщика не описан в стандарте [1], а разработка и анализ различных алгоритмов работы планировщика выходят за рамки статьи. Поэтому авторами не учитывается влияние МС друг на друга и рассматривается только одна пара БС — МС. Кроме этого, учитывается только нисходящая передача данных, так как количество передаваемых данных от БС к МС существенно больше, чем в обратном направлении (5:1) [9]. В связи с тем, что рассмотрение функционирования планировщика оставлено за пределами данной работы, приняты допущения 3 и 4. Допущение 5 основано на том факте, что составление расписания передач для кадра с номером $k + 1$ осуществляется во время кадра с номером k. Допущение 7 не накладывает каких-либо существенных ограничений на применимость приводимого ниже анализа, который может быть легко расширен для случая с большим количеством уровней потребляемой мощности. Для получения численных значений приняты $P_A = 750$ мВт и $P_S = 50$ мВт [8].

Анализ режима ожидания при пуассоновском входном потоке

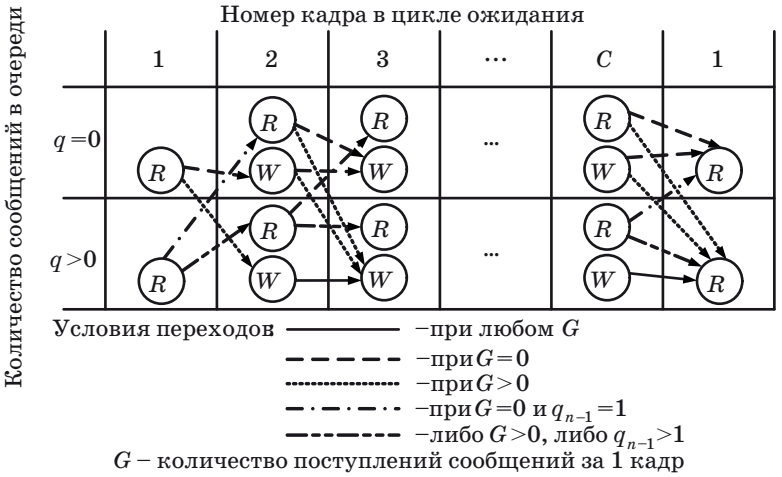
Общий подход к анализу.

Принятая выше модель может быть описана с помощью системы массового обслуживания M/D/1 с перерывами. Однако прямое применение результатов анализа такой модели из известных работ [10, 11] невозможно из-за того, что в нашем случае длительности перерывов (периодов ожидания) имеют различные распределения и зависят между собой.

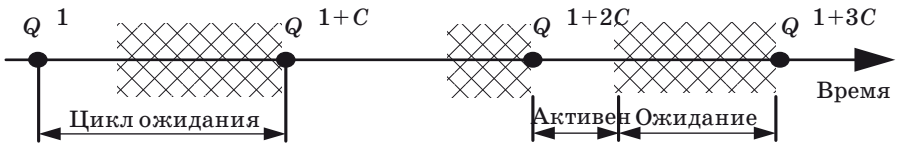
В рамках анализа средней задержки передачи сообщений и энергозатрат МС в условиях функционирования системы в режиме ожидания используются следующие обозначения: C; L; k — номер цикла ожидания; λ — интенсивность входного потока, сообщений/кадр. Для упрощения описания анализа предполагается, что $L = T = 1$.

Рассмотрим случайный процесс (Q^t, M^t), где Q^t — количество сообщений в очереди на БС в начале кадра с номером t; M^t — состояние МС в начале кадра с номером t, которое может принимать два значения: R — активное состояние и W — состояние ожидания. На рис. 2 изображены возможные переходы для данного процесса во время цикла ожидания.

Наблюдать случайный процесс будем в точках начала циклов ожидания (рис. 3). Особенность данных точек состоит в том, что в эти моменты времени система характеризуется только одним параметром — Q^t, потому что в начале цикла



■ ***Рис. 2.*** *Возможные переходы*



■ ***Рис. 3.*** *Формирование вложенной цепи Маркова*

ожидания МС всегда переключается в активное состояние. Отметим, что значение Q^{1+nC} зависит только от $Q^{1+(n-1)C}$, поэтому справедливо следующее утверждение.

Утверждение 1. Последовательность значений Q^1, Q^{1+C}, Q^{1+2C} образует вложенную цепь Маркова.

С ростом интенсивности длительность периодов ожидания уменьшается, и в предельном случае функционирование системы можно рассматривать как функционирование системы M/D/1, которая является устойчивой при любом $\lambda < 1$. Поэтому справедливо следующее утверждение.

Утверждение 2. Определенная выше цепь Маркова имеет стационарное распределение при $\lambda < 1$.

Строгое доказательство последнего утверждения может быть получено из результатов работы [12].

Введем следующее обозначение для переходных вероятностей:

$$p_{ij}^{C} \stackrel{\Delta}{=} \Pr\{Q^{1+nC} = j \,|\, Q^{1+(n-1)C} = i\}.$$

Нижний индекс указывает, как изменилось количество сообщений в очереди на БС, а верхний — за какое количество кадров произошло данное изменение. Расчет переходных вероятностей p_{ij}^{C} осуществляется по следующему рекуррентному выражению:

$$p_{ij}^{C} = \begin{cases} A_j^C, \ i = 0; \\ 0, \ i \geq C, \ j-(i-C) < 0; \\ A_{j-i+C}^C, \ i \geq C, \ j-(i-C) \geq 0; \\ \sum\limits_{k=0}^{j+C-i} p_{kj}^{C-i}, \ 0 < i < C, \end{cases}$$

где A_j^C — вероятность того, что за C кадров поступит ровно j новых сообщений. Для случая с пуассоновским входным потоком $A_j^C = \dfrac{(\lambda C)^j}{j!} e^{-\lambda C}$.

Граничное условие

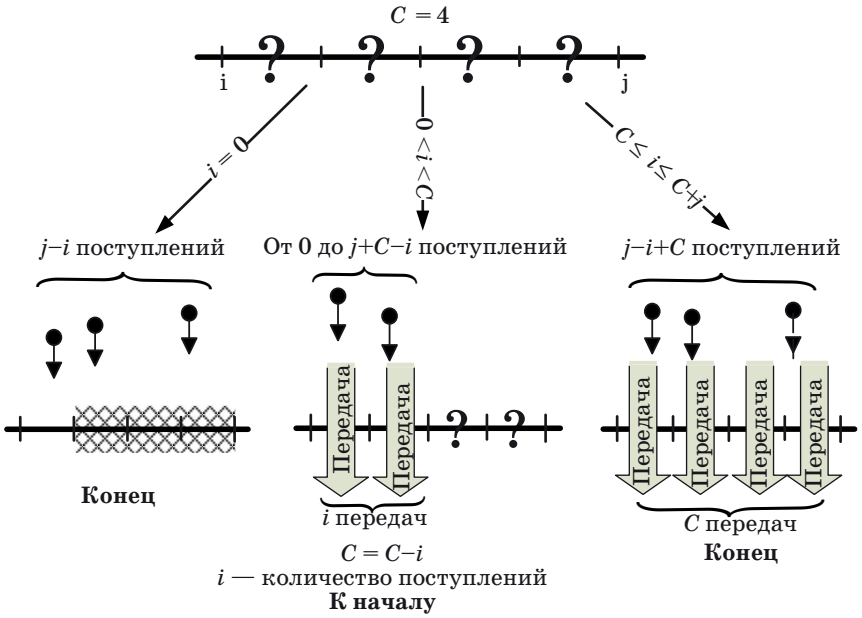
$$p_{ij}^{1} = \begin{cases} A_j^1, \ i = 0; \\ 0, \ i \geq C, \ j-(i-C) < 0; \\ A_{j-i+C}^1, \ i \geq C, \ j-(i-C) \geq 0. \end{cases}$$

Пример одной итерации расчета переходных вероятностей p_{ij}^{C} при $C = 4$ приводится на рис. 4.

Введем определение матрицы **P** вероятностей переходов, состоящей из элементов p_{ij}^{C}:

$$\mathbf{P} = [p_{ij}^{C}].$$

Отметим, что матрица **P** имеет следующий вид:



■ ***Рис. 4.*** *Пример расчета переходных вероятностей*

$$\mathbf{P} = \begin{vmatrix} p_{00}^{C} & p_{01}^{C} & p_{02}^{C} & \cdots \\ p_{10}^{C} & p_{11}^{C} & p_{12}^{C} & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ p_{C0}^{C} & p_{C1}^{C} & p_{C2}^{C} & \cdots \\ 0 & p_{C0}^{C} & p_{C1}^{C} & \cdots \\ 0 & 0 & p_{C0}^{C} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}.$$

Используя специфический вид матрицы, можно рассчитать стационарные вероятности нахождения i сообщений в очереди на БС в начале произвольного цикла ожидания, как показано в работе [13].

Расчет средней задержки и энергозатрат.

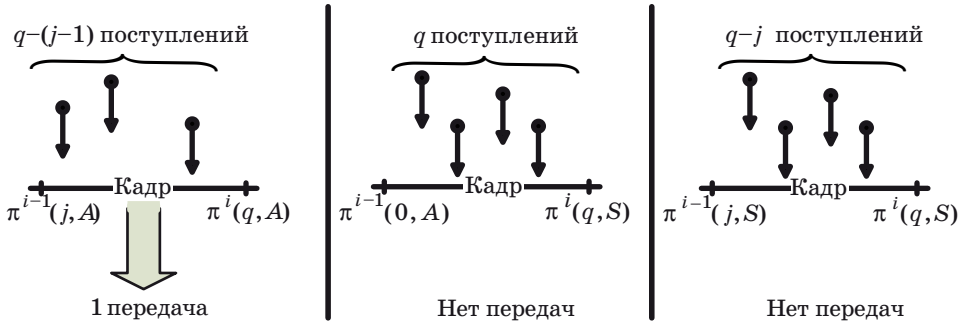
Для расчета средней задержки и величины энергозатрат необходимо знать стационарное распределение количества сообщений в очереди на БС в начале каждого кадра. Выше был приведен механизм расчета стационарного распределения в начале первого кадра цикла ожидания. Соответственно необходимо рассчитать стационарное распределение в начале каждого кадра внутри цикла ожидания. Введем обозначение данных стационарных распределений

$$\pi^{i}(q, m) = \lim_{k\to\infty} \Pr\{Q^{i+C(k-1)} = q, M^{i+C(k-1)} = m\},$$

где $i = [1, C]$.

Возможные переходы за один кадр внутри цикла ожидания изображены на рис. 5. Из рис. 2 и 5 следует, что расчет вероятностей $\pi^{i}(q, m)$ внутри цикла ожидания осуществляется по следующим формулам:

$$\pi^{i}(q, R) = \sum_{j=1}^{q+1} \pi^{i-1}(j, R) A_{q-(j-1)}^{1}; \qquad (1)$$



■ ***Рис. 5.*** *Возможные переходы за один кадр*

$$\pi^i(q,W)=\sum_{j=0}^{q}\pi^{i-1}(j,W)A_{q-j}^1+\pi^{i-1}(0,R)A_q^1, \quad (2)$$

где A_i^1 — вероятность того, что за один кадр поступит i новых заявок.

Выражение (1) применяется для случая, когда МС находится в активном состоянии (т. е. в предыдущем кадре была передача данных). Переход в активное состояние с q сообщениями в очереди на БС возможен только в том случае, когда МС была в активном состоянии в начале предыдущего кадра и в очереди было хотя бы одно сообщение. Выражение (2) используется для случая, когда МС находится или переходит в состояние ожидания. Если в начале предыдущего кадра сообщений в очереди на БС не было, то МС переключится в состояние ожидания.

Используя полученные ранее значения $\pi^i(q, m)$, можно вычислить среднюю длину очереди в начале каждого кадра в рамках цикла ожидания по формуле

$$E[q^i]=\sum_{l=0}^{\infty}l(\pi^i(l,R)+\pi^i(l,W)).$$

Далее используем обобщенную формулу Литтла для вычисления средней задержки передачи данных

$$E[D]=\frac{\sum_{i=1}^{C}E[q^i]}{C\lambda}+\frac{F}{2},$$

где $E[q^i]$ — среднее количество сообщений в очереди на БС в начале i-го кадра цикла ожидания; λ — средняя интенсивность входного потока; F — длительность кадра.

Принимая во внимание допущение 7, вычисление средней величины энергопотребления за цикл ожидания осуществим по формуле

$$E[E]=P_A+\sum_{i=2}^{C}\left(P_S\sum_{j=1}^{i-1}\pi^j(0,R)+P_A\left(1-\sum_{j=1}^{i-1}\pi^j(0,R)\right)\right).$$

Учет влияния изменчивости входного потока

Изменение модели входного потока.

В настоящее время имеется ряд исследований, в которых утверждается, что потоки данных, возникающие в реальных системах передачи данных, не являются пуассоновскими [14]. Существует множество моделей потоков данных, которые рекомендованы к использованию при изучении современных систем передачи данных (например, [15—17]). Однако данные модели отличаются сложностью построения и анализа. Для изучения влияния изменчивости входного потока на рассматриваемые характеристики системы предлагается использовать упрощенную модель входного потока. Изменим допущение относительно модели входного потока следующим образом.

Допущение 6′. В начале каждого кадра определяется состояние источника сообщений. С вероятностью P_{ON} источник находится в состоянии ON, а с вероятностью P_{OFF} — в состоянии OFF. Число сообщений, поступивших в состоянии ON, является случайной величиной, распределенной по закону Пуассона с интенсивностью λ_{ON}. В состоянии OFF поступлений сообщений нет.

Описанную выше модель можно рассматривать как частный случай дважды стохастического пуассоновского процесса (ДСПП). Далее такую модель потока мы будем называть потоком со «всплесками». Используя результаты работы [18], приводимый выше анализ на случай с ДСПП-потоком.

Обобщение анализа.

Рассмотрим модель системы, в которой используется входной поток со случайной интенсивностью. В данном случае алгоритм расчета средней задержки и энергозатрат такой же, как и для случая с пуассоновским входным потоком, кроме расчета вероятности A_i^C. При входном потоке со случайной интенсивностью эта вероятность рассчитывается следующим образом:

$$A_i^C=\sum_{j=1}^{C}P_{\text{OFF}}^jP_{\text{ON}}^{C-j}\binom{C}{j}\frac{(\lambda_{\text{ON}}(C-j))^i}{i!}e^{\lambda_{\text{ON}}(C-j)}.$$

Решение оптимизационной задачи

Из описания алгоритма работы режима ожидания и введенной системы допущений следует справедливость следующего утверждения.

Утверждение 3. Функция средней задержки $f_D(C, 1, 1)$ является монотонно возрастающей при увеличении длительности цикла ожидания C, в то время как функция энергозатрат $f_E(C, 1, 1)$ является монотонно убывающей.

Как следует из данного утверждения, для решения оптимизационной задачи достаточно найти максимальное значение длительности периода ожидания $C_{\max}$, для которого выполняется ограничение из оптимизационной задачи

$$f_D(C_{\max},1,1)\le D_{\max},$$

являющееся оптимальным: $C_{\max}=C_{\text{opt}}$.

Зависимость оптимальной длительности цикла ожидания от средней интенсивности входного

потока при разных типах потока представлена на рис. 6, *а*. Для построения данного графика было выбрано ограничение на среднюю задержку передачи данных, равное 30 мс. На графике изображены результаты решения оптимизационной задачи, т. е. отмечены длительности циклов ожидания, при которых соблюдается ограничение на среднюю задержку и минимизируются затраты МС.

Вид зависимости оптимальной длительности цикла ожидания от интенсивности входного потока объясняется тем, что в области малых интенсивностей задержка передачи в большей степени определяется длительностью цикла ожидания. Отсюда, продолжительные циклы ожидания недопустимы. С ростом интенсивности потока все большая часть цикла ожидания отводится под прием данных, поэтому длительность цикла ожидания может быть увеличена с соблюдением ограничения на среднюю задержку. Однако с дальнейшим ростом интенсивности входного потока система приближается к точке насыщения, поэтому необходимо уменьшать длительность цикла ожидания.

Также на графике рис. 6, *а* отражены результаты для случая ДСПП-потока с различными вероятностями P_{ON}. При одинаковой средней продолжительности состояний ON и OFF источника сообщений сохраняется такой же вид рассматриваемой зависимости, как и при пуассоновском входном потоке. Однако из-за наличия «всплесков» наблюдаются меньшие значения длительности цикла ожидания в сравнении с пуассоновским входным потоком. Кроме этого, с уменьшением длительности периода ON (и, соответственно, с увеличением длительности OFF периода) изменяется характер зависимости. Данная зависимость становится монотонно убывающей.

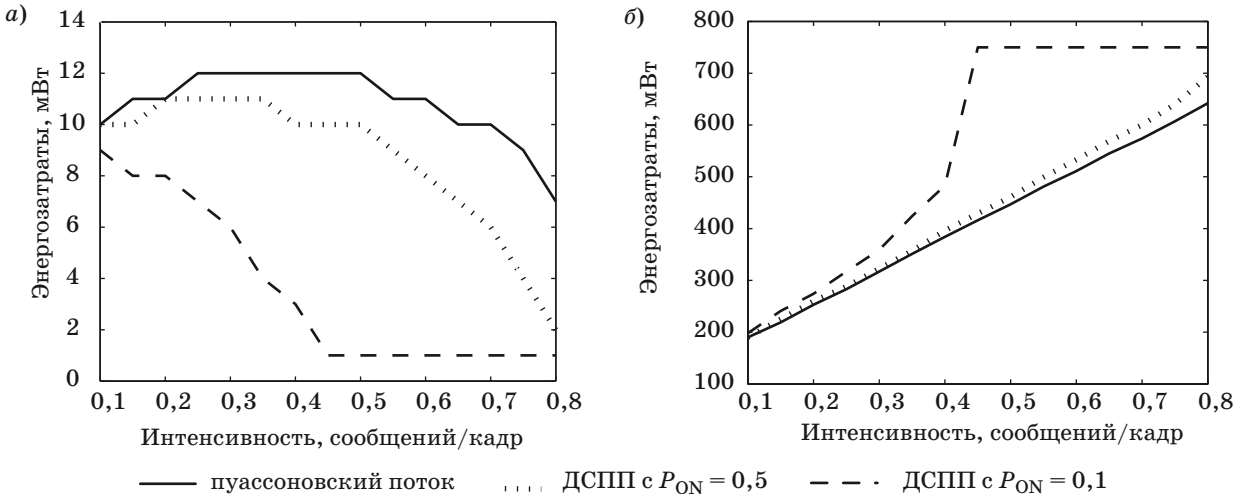
Зависимость энергозатрат от средней интенсивности входного потока изображена на рис. 6, *б*. Здесь же изображены зависимости для различных моделей входного потока. В случае пуассоновского входного потока энергозатраты линейно увеличиваются с ростом средней интенсивности. Для случая с ДСПП с $P_{ON} = 0{,}5$ существует также линейная зависимость на большем диапазоне интенсивностей, за исключением области высоких интенсивностей, так как там наблюдается экспоненциальный рост энергозатрат. Кроме этого, экспоненциальная зависимость наблюдается и для случая с ДСПП с $P_{ON} = 0{,}1$.

Заключение

На основе приведенных графиков и рассуждений можно сделать следующий вывод.

Если выбирать длительность цикла ожидания исходя только из значения средней интенсивности и предполагая, что входной поток представляет собой пуассоновский процесс, в случае потока со «всплесками» ограничения на среднюю задержку передачи данных будут нарушены, хотя величина энергозатрат МС будет минимальной. Если предполагать, что входной поток имеет сильно выраженные «всплески», и выбирать значение цикла ожидания исходя из этого, ограничение на среднюю задержку передачи данных будет соблюдено, однако величина энергозатрат будет достаточно высокой.

При корректном выборе параметров режима ожидания можно обеспечить требуемое качество обслуживания как для случая приема потока со «всплесками», так и для пуассоновского потока.



■ ***Рис. 6.*** *Зависимость оптимальной длины цикла ожидания* (*а*) *и энергозатрат* (*б*) *от интенсивности потока*

Литература

1. **IEEE** P802.16m/D5. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. Apr. 2010. — 925 p.
2. **Вишневский В., Портной С., Шахнович И.** Энциклопедия WiMAX. Путь к 4G. — М.: Техносфера, 2009. — 472 с.
3. **Park Y., Hwang G.** Performance modelling and analysis of the sleep mode in IEEE 802.16e WMAN// Proc. of the 65th IEEE Vehicular Technology Conf. 2007. P. 2801–2806.
4. **Park Y., Hwang G.** An efficient power saving mechanism for delay-guaranteed services//IEEE 802.16e. IEICE Transactions on Communications. 2009. Vol. 1. P. 277–287.
5. **Alouf S., Altman E., Azad A.** Analysis of an M/G/1 queue with repeated inhomogeneous vacations with application to IEEE 802.16e power saving mechanism// Proc. of the 5th Intern. Conf. on Quantitative Evaluation of Systems. 2008. P. 27–36.
6. **Mohammad N., Nejatian P., Nayebi M.** Evaluating the effect of non-Poisson traffic patterns on power consumption of sleep mode in the IEEE 802.16e MAC// Proc. of the Intern. Conf. on Wireless and Optical Communications Networks. 2007. P. 1–5.
7. **DeTurck K.** et al. Performance of the IEEE 802.16e sleep mode mechanism in the presence of bidirectional traffic//Proc. of the Intern. Workshop on Green Communications. 2009. P. 1–5.
8. **Baek S., Son J., Choi B.** Performance analysis of sleep mode operation for IEEE 802.16m advanced WMAN// Proc. of the IEEE Intern. Conf. on Communications Workshops. 2009. P. 1–4.
9. **Flament M.** et al. An approach to 4th Generation Wireless Infrastructures: Scenarios and Key Research Issues//Proc. VTC'99. Houston, TX, May 1999. P. 16–20.
10. **Клейнрок Л.** Теория массового обслуживания. — М.: Машиностроение, 1979. — 432 с.
11. **Бертсекас Д., Галлагер Р.** Сети передачи данных. — М.: Мир, 1989. — 544 с.
12. **Baccelli F., Foss S.** On the saturation rule for the stability of queues//J. Appl. Prob. 1995. N 32. P. 494–507.
13. **Neuts M. F.** Structured Stochastic Matrices Of M/G/1 Type And Their Applications. — CRC Press, 1989. — 532 p.
14. **Ahson S.** et al. WiMAX Technologies, Performance Analysis, and Qos. — CRC Press, 2008. — 296 p.
15. **Traffic** Model for 802.16 TG3 MAC-PHY Simulations// IEEE 802.16.3c-01/30r1. Mar. 2001. — 27 p.
16. **Mean** Traffic Bit Rate with ON-SID Modeling of VoIP Traffic//IEEE C802.16m-07/123, July 2007. — 6 p.
17. **WiMAX** System Evaluation Methodology. Version 2.1, July 2008. — 209 p.
18. **Хименко В.** Характеристики типа «превышений уровней» для случайных точечных процессов // Радиотехника и электроника. 2000. Т. 45. № 4. С. 436–443.

УДК 004.07

О ПРОПУСКНОЙ СПОСОБНОСТИ БЕСПРОВОДНЫХ МНОГОКАНАЛЬНЫХ ОДНОРАНГОВЫХ СЕТЕЙ С МЕСТНЫМ ПЛАНИРОВАНИЕМ ЧАСТОТНОГО РАЗДЕЛЕНИЯ КАНАЛОВ[1]

Й. Эльснер,
научный сотрудник
Р. Танбурги,
научный сотрудник
Ф. Йондраль,
доктор техн. наук, профессор, директор института
Технологический институт г. Карлсруэ, Германия

Проведен анализ пропускной способности беспроводных многоканальных одноранговых сетей, ограниченных воздействием помех, с местным планированием частотного разделения каналов. Выведены верхняя и нижняя границы вероятностей прерывания связи и пропускной способности в данной модели системы. Выполнено сравнение с многоканальными сетями без местного планирования частотного разделения каналов.

Ключевые слова *— многоканальные одноранговые сети, частотное разделение каналов, пропускная способность.*

Введение

Исследования плотных, беспроводных сетей, ограниченных помехами, вызывают в последнее время повышенный научный интерес. С точки зрения теории информации, использовать широкую или даже бесконечную полосу пропускания для передачи данных по каналу имеет смысл только при отсутствии помех. Помехи, т. е. необходимость сосуществовать с другими передатчиками, являются основной причиной ограничения полосы пропускания системы. Всевозможные технические проблемы также усложняют дизайн широкополосных беспроводных систем. Из-за таких требований, как необходимость большого динамического диапазона или низкой потребляемой мощности, пропускная полоса на физическом уровне гораздо уже, чем диапазон возможных центральных частот передатчика. Таким образом, возникает вопрос пропускной способности таких систем, а точнее, требуется ответ на вопросы, «как с точки зрения теории информации выбрать ширину полосы пропускания и сколько передатчиков в сети могут одновременно работать?». В статье исследуются эти вопросы в модели Вебера [1].

Более ранними, связанными с данной проблемой, являются статьи [2–6], из которых большинство написано в соавторстве с Вебером. В работах [2, 3] сравниваются методы расширения спектра в одноранговых сетях — псевдослучайная перестройка рабочей частоты и расширение спектра методом прямой последовательности — под условием эквивалентной мощности. По результатам анализа, системы псевдослучайной перестройки рабочей частоты эффективнее, чем расширение спектра методом прямой последовательности в отношении пропускной способности. Авторы [1] в другой статье [4] рассматривают проблему разделения операционной пропускной полосы в целях максимизации спектральной эффективности на квадратный метр. Пропускная способность в сетях с компенсацией помех выведена в работе [5]. Мы же рассматриваем разделение операционной пропускной полосы таких сетей, которые могут местно планировать трансмиссии методом частотного разделения каналов, при этом анализ наш основывается на вышеуказанных результатах Вебера и др.

[1] Переведено с английского оригинала, публикуется с исправлениями. Реф. в: Proc. of the Intern. Congress on Ultra Modern Telecommunications and Control Systems, Moscow, 2010. © IEEE, 2010.

Модель системы

Беспроводная сеть состоит из узлов $\{X_i\}$, распределенных на плоскости согласно однородному пуассоновскому точечному процессу (ПТП) Π интенсивностью λ, где $X_i \in \mathbb{R}^2$ обозначает позиции передатчиков, которые вызывают помехи. Операционная полоса пропускания всей сети B разделена на M ортогональных каналов с системной полосой пропускания $B_m = \frac{B}{M}$. Спектральная плотность мощности окружающего шума — N_0. Коэффициент затухания между двумя точками в плоскости на расстоянии d дан как $d^{-\alpha}$, $\alpha > 2$. Согласно теореме Сливняка [ср. 7, с. 121], статистика Π не меняется при добавлении измерительной пары приемника и передатчика в каждом канале. Каналы являются симметричными, поэтому достаточно рассмотреть одну измерительную пару: приемник, который находится в начале координат, и передатчик, расположенный на расстоянии r метров от него. Каждый передатчик, как измерительный, так и источники помех, работает с мощностью ρ, и, следовательно, спектральная плотность мощности в каждом канале составляет $\frac{\rho}{B_m} = \frac{\rho M}{B}$.

Рабочая характеристика приемника зависит от помех в канале, порожденных узлами, работающими в том же канале. Так как ПТП однородный, рабочая характеристика приемника также отражает рабочую характеристику всей сети.

Мы предполагаем, что сетевой протокол способен ортогонализировать передачи всех соседних передатчиков вокруг приемника в радиусе ортогонализации r_o. Трансляции в одном и том же районе осуществляются в различных каналах. Таким образом вокруг приемника создается зона, свободная от источников помех.

Чем больше существует каналов, тем меньше влияние помех, так как это позволяет ортогонализировать большее количество соседних узлов. Тем самым сокращается влияние помех в канале, поскольку деятельность передатчиков делится на M каналов. С другой стороны, это приводит к сокращению спектра для связи в одном канале, следовательно, вероятность прерывания связи повышается. В следующих разделах мы определим оптимальное количество каналов.

Пропускная способность без местного планирования

Во-первых, выводится соотношение между шириной полосы пропускания системы, шириной операционной полосы пропускания и пропускной способностью для сетей без местного планирования каналов, которые должны поддерживать минимальную скорость передачи R_m от одного узла к другому на максимальном расстоянии r. В этом случае существует оптимальное количество каналов, которое максимизирует пропускную способность независимо от интенсивности λ.

Отношение уровня сигнала к уровню шума и помех (SINR) у измерительного передатчика в одном канале m можно выразить как

$$\mathrm{SINR} = \frac{\rho r^{-\alpha}}{N_0 B_m + \sum\limits_{i \in \Pi_m} \rho |X_i|^{-\alpha}},$$

где Π_m обозначает ПТП с интенсивностью $\lambda_m = \frac{\lambda}{M}$, так как деятельность передатчиков делится на M каналов. В канале m вероятность прерывания связи q_m при скорости передачи R_m составляет

$$q_m(\lambda_m) = P\{B_m \log_2(1+\mathrm{SINR}) \le R_m\} = P\left\{\underbrace{\mathrm{SINR}^{-1}}_{Y_m} > \left(2^{\frac{R_m}{B_m}} - 1\right)^{-1} = \frac{1}{\beta}\right\}.$$

С учетом симметричности двухмерный ПТП Π_m можно отобразить как одномерный ПТП с интенсивностью 1 [1, 8] следующим образом:

$$Y_m = \frac{N_0 B_m}{\rho r^{-\alpha}} + \left(\pi r^2 \lambda_m\right)^{\frac{\alpha}{2}} \underbrace{\sum_{i \in \Pi_1} T_i^{-\frac{\alpha}{2}}}_{Z_\alpha},$$

где T_i — расстояние передатчика i, вызывающего помехи, от начала координат. Для q_m следует

$$q_m(\lambda_m, R_m) = P\left\{Z_\alpha > \left(\pi r^2 \lambda_m\right)^{-\frac{\alpha}{2}} \underbrace{\left(\left(2^{\frac{R_m}{B_m}} - 1\right)^{-1} - \frac{N_0 B_m}{\rho r^{-\alpha}}\right)}_{\theta_m}\right\}.$$

Обозначим комплемент кумулятивной функции распределения Z_α как $\bar{F}_{z_\alpha}$, тогда

$$q_m(\lambda_m) = \bar{F}_{z_\alpha}\left(\left(\pi r^2 \lambda_m\right)^{-\frac{\alpha}{2}} \theta_m\right). \qquad (1)$$

Оптимальная пропускная полоса каждого канала выводится путем минимизации вероятностей прерывания связи:

$$M_{\text{opt}} = \arg\min_{M} q_m(\lambda_m, R_m).$$

С учетом уравнения (1), $\lambda_m = \frac{\lambda}{M}$ и $B_m = \frac{B}{M}$ следует

$$M_{\text{opt}} = \arg\min_{M}\left(\bar{F}_{z_\alpha}\left(\left(\pi r^2 \lambda_m\right)^{-\frac{\alpha}{2}}\theta_m\right)\right) =$$

$$= \arg\max_{M}\left(M^{\frac{\alpha}{2}}\left(\left(2^{\frac{R_m}{B_m}}-1\right)^{-1} - \underbrace{\frac{N_0 B}{\rho r^{-\alpha} M}}_{N_0/E_b}\right)\right). \qquad (2)$$

Эту проблему оптимизации можно решить в замкнутой форме в случае сети, ограниченной помехами: $\frac{E_b}{N_0} \to \infty$ [4]:

$$M_{\text{opt}} = \frac{B}{R_m}\underbrace{\log_2 \exp\left(\frac{\alpha}{2} + W\left(-\frac{\alpha}{2}\exp\left(-\frac{\alpha}{2}\right)\right)\right)}_{b_{\text{opt}}}, \qquad (3)$$

где $W(\cdot)$ обозначает основную отрасль W-функции Ламберта. Согласно уравнению (3), оптимальное разделение пропускной полосы зависит от α, желаемой скорости передачи информации R_m и ширины операционной полосы B. Оптимальное разделение соответствует порогу SINR β с величиной

$$\beta_{\text{opt}} = 2^{b_{\text{opt}}} - 1$$

и спектральной эффективностью b_{opt}.

Пропускная способность одного канала определяется как [1]

$$c_m(q_m) = \lambda_m(1 - q_m),$$

где функция $q_m^{-1}(\varepsilon) = \lambda_m$ — обратная функция вероятности прерывания связи; $q_m^{-1}(\varepsilon)$ дает пространственную интенсивность трансмиссии λ_m при вероятности прерывания связи $\varepsilon \in [0, 1]$, и эта интенсивность умножается на вероятность удачной трансмиссии.

Следовательно, пропускная способность — мера удачных трансмиссий в пространстве.

В многоканальных сетях общая пропускная способность дается как

$$c(\varepsilon) = \sum_{m=1}^{M} c_m(q_m).$$

Таким образом, из (1) и (3) следует пропускная способность

$$c(\varepsilon) = \frac{\bar{F}_{z_\alpha}^{-1}(\varepsilon)^{-\frac{2}{\alpha}}}{\pi r^2}(1-\varepsilon)\beta_{\text{opt}}^{-\frac{2}{\alpha}}\underbrace{\frac{B}{R_m}}_{M_{\text{opt}}} b_{\text{opt}}. \qquad (4)$$

Пропускная способность растет прямо пропорционально ширине операционной пропускной полосы B. В многоканальных сетях отношение $\frac{R_m}{B}$ будет низким, а количество каналов, соответственно, высоким. Как правило, пропускную способность (4) невозможно выразить в замкнутой форме. Исключением является случай $\alpha = 4$, где пропускная способность с нотацией $Q(z) = Q\{Z \le z\}$ и $Z \sim N(0, 1)$ равна [1]

$$c(\varepsilon) = \frac{\sqrt{2/\pi}Q^{-1}\left((1+\varepsilon)/2\right)}{\pi r^2\sqrt{\beta_{\text{opt}}}}(1-\varepsilon)M_{\text{opt}}.$$

Этот случай будет использован в качестве сравнения.

Пропускная способность с местным планированием

Возможность ортогонализации.

Важный вопрос, которым нельзя не задаться, возможность ортогонализации N соседних узлов, если существуют $M = N + 1$ каналов. Аналогичная задача — раскраска вершин бесконечного случайного графа цветами в количестве M. Рабочая характеристика измерительного приемника точно отражает рабочую характеристику всех узлов, только если ортогонализация возможна во всех вершинах графа с высокой степенью вероятности. Согласно теореме Брукса [9], ортогонализация (раскраска) графа с множеством вершин $\{X_i\}$ и M каналами (цветами) возможна, если максимальное число соседних узлов $\max\{N_i\}$ (максимальная степень графа) меньше, чем M. Очевидно, что такое может быть только если число узлов конечно. Следовательно, ПТП обусловлен на K узлов на площади $A \subset \mathbb{R}^2$, где $\frac{K}{A} = \lambda$. Если дан ПТП с интенсивностью λ, то число соседних узлов N_i каждого узла X_i на расстоянии r_o независимо и одинаково распределено по закону Пуассона с $\lambda_n = \pi r_o^2 \lambda$. Достаточным для возможности ортогонализации с вероятностью большей, чем $1 - \varepsilon_o$, является условие, что максимум множества пуассоновских случайных величин, которое описывает число соседних узлов, не превышает число каналов (цветов):

$$P\left\{\max\{N_1, N_2, \ldots, N_K\} \le M-1\right\} > 1-\varepsilon_o.$$

Пользуясь независимостью случайных величин, это условие можно выразить как

$$1-\varepsilon_o(M) < \left(\sum_{i=0}^{M-1} \exp(-\lambda_n)\frac{\lambda_n^i}{i!}\right)^K = \phi(M, \lambda_n)^K, \quad (5)$$

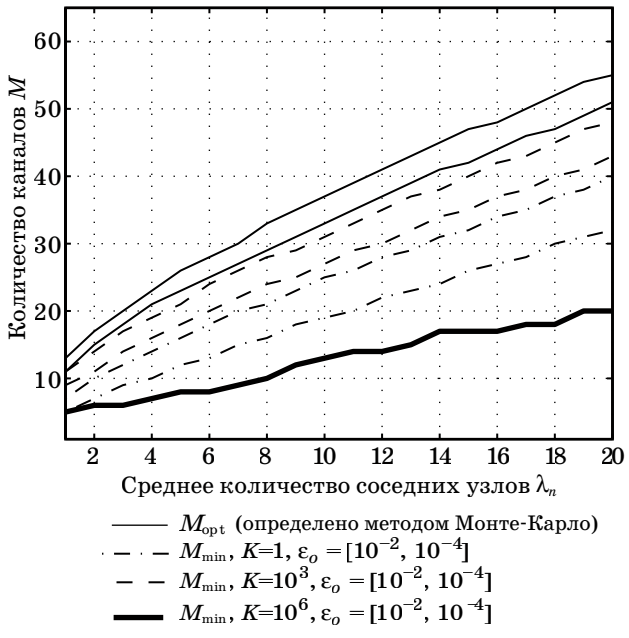
где $\phi(M, \lambda_n) = \frac{\Gamma(M, \lambda_n)}{\Gamma(M)}$ — регуляризованная гамма-функция. Если M растет, $\varepsilon_o(M)$ быстро сводится к 0. Верхней границей $\varepsilon_o(M)$ с нотацией $\exp(x) = \sum_{i=0}^{M-1} \frac{x^i}{i!} + R_M(x)$ является

$$\varepsilon_o(M) = 1 - \left(\sum_{i=0}^{M-1} \exp(-\lambda_n)\frac{\lambda_n^i}{i!}\right)^K =$$

$$= 1 - \left(1 - \exp(-\lambda_n) R_M(\lambda_n)\right)^K \le$$

$$\le 1 - \left(1 - \exp(-\lambda_n)\frac{2\lambda_n^M}{M!}\right)^K$$

для $M \ge 2\lambda_n - 2$ (так как $R_M(x) \le \frac{2|x|^M}{M!}$, доказательство выполняется способом ограничения остаточных членов в сумме ряда).



■ ***Рис. 1.*** *Число каналов, необходимое для ортогонализации в сети с K узлами с вероятностью ε_o, и число каналов, минимизирующее прерывание с местным планированием каналов: $R_m/B = 0{,}1$; $\lambda_n = 5$; $\alpha = 4$; $r = 10$*

Если даны A, λ, $\varepsilon_o > 0$ и r_o, необходимое количество каналов M для ортогонализации в радиусе r_o следует из уравнения (5)

$$M \ge \phi^{-1}\left((1-\varepsilon_o)^{\frac{1}{K}}, \lambda_n\right).$$

Необходимое количество каналов для разных K и $\varepsilon_o > 0$ показано на рис. 1 (каждая нижняя кривая соответствует $\varepsilon_o = 10^{-2}$, а верхняя — $\varepsilon_o = 10^{-4}$). Даже для больших K и маленьких ε_o необходимое количество каналов растет медленно. Это одно из свойств распределения максимума множества независимых пуассоновских случайных величин, рассмотренное Бриггсом [10].

Вероятность прерывания связи и пропускная способность.

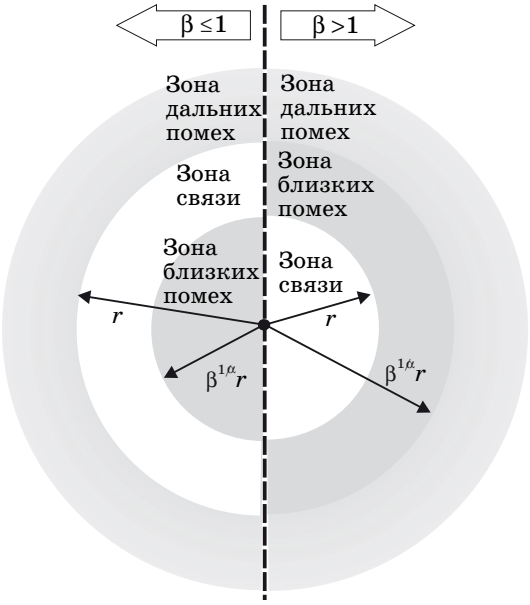
Теперь допустим, что в сети используется M каналов, а сетевой протокол может координировать трансмиссии соседних узлов таким образом, что они по возможности становятся ортогональными. Вероятность возможной ортогонализации $1 - \varepsilon_o$ в сети с K узлами и M каналами описана в (5). Только в том случае, если эта вероятность высокая, возможность местной ортогонализации является представительной мерой рабочей характеристики всей сети. Эта вероятность равна тому, что в радиусе r_o от измерительного приемника имеется не больше чем $M - 1$ узлов:

$$p_o = \sum_{i=0}^{M-1} \exp(-\lambda_n)\frac{\lambda_n^i}{i!}.$$

В последующем предполагается, что уровень управления доступом к среде (MAC) может устранять $M - 1$ близких передатчиков, вызывающих помехи, на расстоянии r способом местного частотного разделения каналов. Таким образом, $r_o = r$. Это предположение оправданно, так как если узлы могут связываться напрямую, они могут эффективно планировать трансмиссии.

Как и в случае без местного планирования, для такого интерференционного поля в замкнутой форме нет решений. Поэтому выводят верхнюю и нижнюю границы вероятностей прерывания связи.

Случай прерывания можно проанализировать, рассматривая зону связи, зону близких помех и зону дальних помех. Зоны связи и близких помех — диски радиусом r и $r_s = \beta^{\frac{1}{\alpha}} r$. Зона дальних помех — район вне зоны близких помех. Передачи в зоне близких помех напрямую вызывают прерывание. Передачи в дальней зоне могут



■ ***Рис. 2.*** *Зоны для* $\beta \le 1$ *и* $\beta > 1$

вызывать прерывание, если мощность помех превышает $\frac{1}{\beta}$.

Следует различать два случая (рис. 2). Если $\beta \le 1$, зона связи равна или больше зоны близких помех; если $\beta > 1$, зона связи меньше, чем зона близких помех. Для обоих случаев выводят границы вероятности прерывания для данного M и соответствующую пропускную способность сети.

Низкая спектральная эффективность, $\beta \le 1$, $\frac{R_m}{B} M \le 1$.

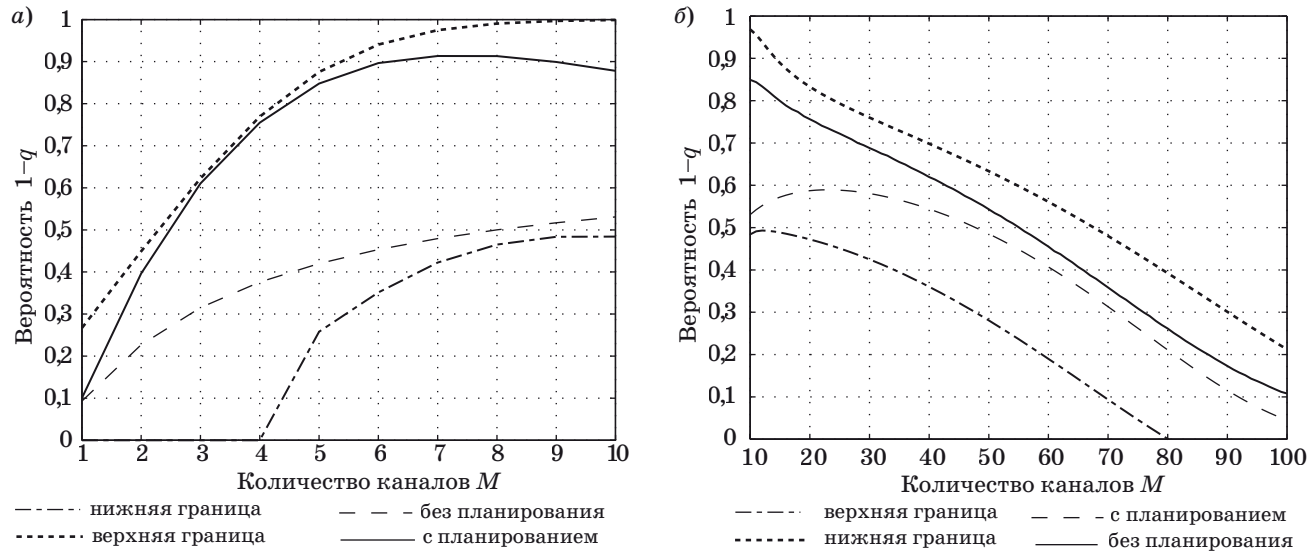
Из $2^{\frac{R_m}{B}M} - 1 = \beta \le 1$ непосредственно следует $\frac{R_m}{B} M \le 1$. В этом случае зона связи больше, чем зона близких помех. Вероятность прерывания связи $q(\lambda)$ состоит из прерывания связи по двум причинам: невозможной ортогонализации вокруг измерительного приемника и помех дальних передатчиков. Эквивалентно вероятность удачной трансмиссии

$$1 - q(\lambda) = p_o(\lambda) P\left\{Y_r(\lambda) \le \beta^{-1}\right\} + \left(1 - p_o(\lambda)\right) P\{\mathrm{F}\} P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\}. \quad (6)$$

Здесь $Y_r(\lambda)$ — количество помех, созданное узлами на расстоянии r или больше. Если ортогонализация у измерительного приемника не удалась, то $P\{F\}$ является вероятностью случая F «нет неортогонализированного узла в зоне близких помех» и $P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\}$ является вероятностью случая «величина помех из дальней зоны помех не больше $\frac{1}{\beta}$ ».

Нижняя граница. Если опустить второе слагаемое в (6), нижняя граница вероятности удачной трансмиссии

$$1 - q(\lambda) > p_o(\lambda) P\left\{Y_r(\lambda) \le \beta^{-1}\right\}. \quad (7)$$



■ ***Рис. 3.*** *Вероятность удачной трансмиссии в случае* $\beta \le 1$ *(а) и* $\beta > 1$ *(б):* $R_m/B = 0{,}1$; $\lambda_n = 5$; $\alpha = 4$; $r = 10$

Член $P\{Y_r(\lambda) \le \beta^{-1}\}$ можно дальше ограничить с помощью неравенства Маркова $cP\{X \ge c\} \le E\{X\}$ с результатом

$$P\left\{Y_r(\lambda) \le \beta^{-1}\right\} \ge 1 - \frac{\lambda}{M}\frac{2\pi r^2}{\alpha - 2}\beta.$$

Эта граница не является тесной для малых M. Для $M \to 1$ $p_o = 0$, и, следовательно, (6) заменяется на $P\{\text{нет узла в } r_s\}P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\}$.

Верхняя граница. Верхнюю границу вероятности удачной трансмиссии выводят, предполагая, что измерительный приемник всегда способен ортогонализировать соседние $M - 1$ узлов, и учитывая только зону близких помех. В этом случае трансмиссия удачна, если в зоне близких помех находятся не больше, чем $M - 1$ передатчиков. $P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\}$ можно ограничить следующим образом:

$$P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\} \le \sum_{i=0}^{M-1} \exp(-\lambda_s)\frac{\lambda_s^i}{i!}, \qquad (8)$$

где $\lambda_s = \lambda\pi\left(\beta^{\frac{1}{\alpha}} r\right)^2$. Границы пропускной способности следуют из численного обращения (7) и (8) для λ. Рис. 3, *а* показывает вероятность удачной трансмиссии, определенной методом Монте-Карло, границы (7) и (8), а также точный результат для сетей без местного планирования каналов.

Высокая спектральная эффективность, $\beta > 1$, $\frac{R_m}{B}M > 1$.

В режиме высокой спектральной эффективности зона близких помех больше, чем зона связи. В этом случае причиной прерывания может быть неудачная ортогонализация, присутствие одного или большего количества узлов в кольце с радиусами $r_s = \beta^{\frac{1}{\alpha}} r$ и r, а также величина помех из дальней зоны большая, чем $\frac{1}{\beta}$. Вероятность удачной трансмиссии

$$1 - q(\lambda) = p_o(\lambda)P\{\text{нет узла в кольце}\}P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\},$$

где $P\{\text{нет узла в кольце}\} = \exp\left(-\frac{\lambda}{M}\pi(r_s^2 - r^2)\right)$.

Нижняя граница. Нижнюю границу вероятности удачной трансмиссии выводят с помощью неравенства Маркова для $P\left\{Y_{r_s}(\lambda) \le \beta^{-1}\right\}$:

$$1 - q(\lambda) \ge p_o(\lambda)\left(1 - \frac{\lambda}{M}\frac{2\pi r^2}{\alpha - 2}\beta^{\frac{2}{\alpha}}\right) \times$$

$$\times \exp\left(-\frac{\lambda}{M}\pi(r_s^2 - r^2)\right). \qquad (9)$$

Верхняя граница. Верхнюю границу вероятности удачной трансмиссии выводят с $p_o(\lambda) = 1$ и опусканием помех из дальней зоны помех:

$$1 - q(\lambda) \le \exp\left(-\frac{\lambda}{M}\pi(r_s^2 - r^2)\right). \qquad (10)$$

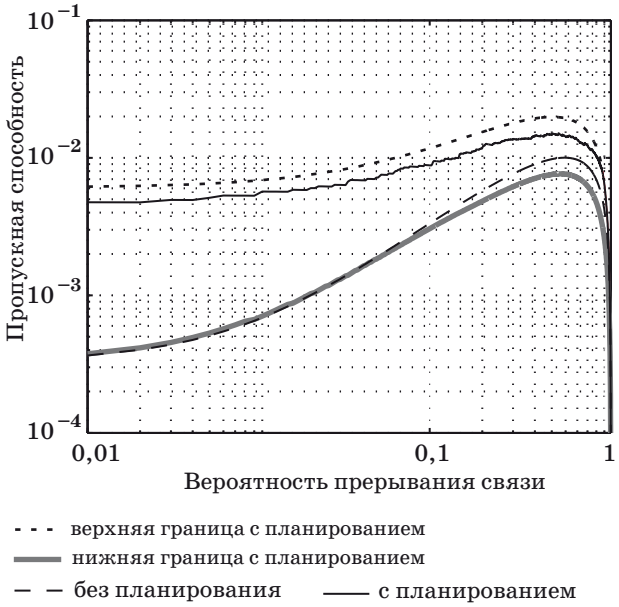
Вероятность удачной трансмиссии, определенной методом Монте-Карло, границы (9) и (10), а также точный результат для сетей без местного планирования каналов показаны на рис. 3, *б*.

Обсуждение результатов.

Для определенного маленького количества соседних узлов λ_n местное планирование ведет к значительному увеличению вероятности удачной трансмиссии, особенно если спектральная эффективность связей низкая (см. рис. 3).

Близкие передатчики, которые вызывают преимущественную часть помех, ортогонализованы, и их не требуется избегать способом случайного выбора канала. В режиме высокой спектральной эффективности увеличение вероятности удачной трансмиссии быстро уменьшается, так как передатчики с доминирующими помехами находятся вне зоны связи.

Пропускная способность сетей с местным планированием и без местного планирования представлена на рис. 4. В случае «без местного планиро-



■ ***Рис. 4.*** *Пропускная способность в случае* $K = 10^3$; $\varepsilon_o = 10^{-2}$; $R_m/B = 0{,}1$; $\alpha = 4$; $r = 10$

вания» количество каналов выбрано согласно (2). В случае «с местным планированием» оптимальное количество каналов M_{opt} всегда меньше, чем число каналов, необходимое для ортогонализации (см. рис. 1). Соответственно, минимальное количество каналов для возможной ортогонализации было выбрано для получения результатов, показанных на рис. 4. Высокая допустимая вероятность прерывания связи ведет к большому количеству каналов и отсюда к высокой спектральной эффективности. Важным фактом является то, что оптимальное количество каналов M_{opt} с местным планированием зависит от интенсивности λ, что не верно в отсутствие местного планирования. Общий коэффициент усиления пропускной способности в данном случае составил от 1,35 до 13, в зависимости от допустимой вероятности прерывания связи.

Заключение

Как видно из анализа, очень полезно использовать местное планирование каналов в одноранговых сетях. Местное планирование способом частотного разделения каналов ведет к значительному росту пропускной способности и является менее сложным, чем, например, компенсация помех каждого приемника. Эффективный MAC для многоканальных одноранговых сетей должен включать в себя динамическое планирование каналов и адаптирование спектральных эффективностей.

Интересное направление дальнейшего исследования — расширение результатов на каналы с фединго м, например по результатам Вебера [1].

Литература

1. **Weber S., Andrews J., Jindal N.** An overview of the transmission capacity of wireless networks//IEEE Transactions on Communications. Dec. 2010. Vol. 58. N 12.
2. **Weber S., Yang X., Andrews J., de Veciana G.** Transmission capacity of wireless ad hoc networks with outage //IEEE Transactions on Information Theory. Dec. 2005. Vol. 51. N 12. P. 4091–4102.
3. **Andrews J., Weber M., Haenggi S.** Ad Hoc Networks: To Spread or Not to Spread?//EEE Communications Magazine. Dec. 2007. Vol. 45. N 12. P. 84–91.
4. **Jindal N., Andrews J., Weber S.** Bandwidth partitioning in decentralized wireless networks//IEEE Transactions on Wireless Communications. Dec. 2008. Vol. 7. N 12. P. 5408–5419.
5. **Weber S., Andrews J., Yang X., de Veciana G.** Transmission capacity of wireless ad hoc networks with successive interference cancellation//IEEE Transactions on Information Theory. Aug. 2007. Vol. 53. N 8. P. 2799–2814.
6. **Hasan A., Andrews J.** The guard zone in wireless ad hoc networks//IEEE Transactions on Wireless Communications. Mar. 2007. Vol. 6. N 3. P. 897–906.
7. **Brooks R.** On colouring the nodes of a network// Mathematical Proc. of the Cambridge Philosophical Society. Apr. 1941. Vol. 37. N 2. P. 194–197.
8. **Stoyan D., Kendall W., Mecke J.** Stochastic geometry and its applications. 2nd ed. — N. Y.: Wiley, 1995.
9. **Haenggi M.** On distances in uniformly random networks//IEEE Transactions on Information Theory. Oct. 2005. Vol. 51. N 10. P. 3584–3586.
10. **Briggs K., Song L., Prellberg T.** A note on the distribution of the maximum of a set of poisson random variables: preprint. Mar. 2009. http://arxiv.org/abs/0903.4373v2 (дата обращения: 26.03.2009).

УДК 681.326.74

СИНТЕЗ ТРИСИНГУЛЯРНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Л. А. Мироновский,
доктор техн. наук, профессор
И. Р. Курмаев,
аспирант
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Выделен класс линейных стационарных систем с тремя группами кратных ганкелевых сингулярных чисел. Для таких систем, названных трисингулярными, установлен вид сбалансированного представления и найдены канонические структурные реализации. Разработан алгоритм декомпозиции трисингулярной передаточной функции на фазовращательные слагаемые и алгоритм синтеза систем с заданными ганкелевыми сингулярными числами и характеристическим полиномом.

Ключевые слова *— линейные стационарные динамические системы, ганкелевы сингулярные числа, грамианы управляемости и наблюдаемости, декомпозиция передаточной функции, фазовращательные подсистемы.*

Введение

Линейные динамические модели широко используются при изучении реальных систем. В теории управления и теории систем изучены различные классы линейных моделей, такие как устойчивые, минимально фазовые, позитивные и др. В настоящей работе исследуется сравнительно новый и недостаточно изученный класс линейных систем с ганкелевыми сингулярными числами (ГСЧ) высокой кратности.

Системы с кратными ГСЧ обладают специальными свойствами и играют важную роль при решении классических задач оптимального управления, идентификации и редукции. Здесь можно упомянуть теорему Неванлины—Пика, расширение Нехари, фазовую декомпозицию Гловера [1–8]. Большинство известных результатов относятся к случаю систем, у которых все ГСЧ совпадают (так называемые моносингулярные системы). Такие системы достаточно часто встречаются в математике и инженерной практике. Например, в радиотехнике находят применение фазовращательные звенья, обладающие постоянной амплитудно-частотной характеристикой. Другой пример — мост Вина—Робинсона, который применяется при построении генераторов синусоидальных колебаний.

В работах [6, 7] исследованы динамические системы с двумя группами одинаковых ГСЧ (так называемые бисингулярные системы). В данной статье исследуется более широкий класс систем, сингулярные числа ганкелева оператора которых образуют три группы кратных чисел (трисингулярные системы). Роль подобных объектов в теории линейных динамических систем аналогична роли операторов с кратными собственными числами в линейной алгебре. Как известно, последние представляют большую трудность для анализа (сложная структура инвариантных и корневых подпространств, наличие жордановых клеток высокого порядка, плохая обусловленность). В то же время наиболее тонкие и глубокие результаты линейной алгебры получены при исследовании именно таких операторов.

Исследование трисингулярных динамических систем (ТДС) требует постановки и решения задач синтеза этих систем. К числу таких задач относятся:

— получение блочно-сбалансированной и фазовой декомпозиций ТДС;

— установление вида передаточной функции ТДС в специальных случаях;

— разработка алгоритмов синтеза ТДС с заданными ГСЧ.

Статья посвящена решению этих задач.

Собственные и сингулярные числа ганкелева оператора

Приведем необходимые сведения о ГСЧ. Рассмотрим устойчивую линейную стационарную систему n-го порядка с одним входом u и одним

выходом y, заданную описанием в пространстве состояний

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} + \mathbf{b}u,\ y = \mathbf{c}\mathbf{X}, \qquad (1)$$

где $\mathbf{A}$ — постоянная $(n \times n)$-матрица; $\mathbf{b}$ и $\mathbf{c}$ — вектор-столбец и вектор-строка.

К числу важных вход-выходных характеристик системы (1), наряду с корнями характеристического полинома, относятся ее ГСЧ. Классический способ их определения основан на рассмотрении грамианов управляемости и наблюдаемости $\mathbf{W}_c$ и $\mathbf{W}_o$, которые могут быть найдены из матричных уравнений Ляпунова

$$\mathbf{A}\mathbf{W}_c + \mathbf{W}_c\mathbf{A}^T + \mathbf{b}\mathbf{b}^T = 0,$$
$$\mathbf{A}^T\mathbf{W}_o + \mathbf{W}_o\mathbf{A} + \mathbf{c}^T\mathbf{c} = 0.$$

Собственные значения произведения грамианов $\mathbf{W}_c\mathbf{W}_o$ не зависят от выбора базиса в пространстве состояний. Если система устойчива, управляема и наблюдаема, то все эти значения вещественны и положительны.

Положительные числа $\sigma_1, ..., \sigma_n$ — арифметические квадратные корни из собственных значений матрицы $\mathbf{W}_c\mathbf{W}_o$ — называются ганкелевыми сингулярными числами системы (1).

При помощи линейной замены переменных описание (1) можно привести к виду, в котором грамианы равны и диагональны:

$$\mathbf{W}_c = \mathbf{W}_o = \mathrm{diag}(\sigma_1, \ldots, \sigma_n), \qquad (2)$$

причем диагональными элементами грамианов служат ГСЧ. Реализация системы (1), удовлетворяющая условию (2), называется сбалансированным представлением. Сбалансированное представление системы единственно, если все ГСЧ различны по величине.

Ганкелевы сингулярные числа скалярной системы могут быть введены с помощью кросс-грамиана $\mathbf{W}_{co}$, определяемого матричным уравнением Ляпунова

$$\mathbf{A}\mathbf{W}_{co} + \mathbf{W}_{co}\mathbf{A} + \mathbf{b}\mathbf{c} = 0. \qquad (3)$$

Собственные значения $s_1, ..., s_n$ кросс-грамиана $\mathbf{W}_{co}$ будем называть ганкелевыми собственными значениями (ГСЗ) системы (1). Они отличаются от ГСЧ только знаками, т. е. имеют место равенства $s_1 = i_1\sigma_1$, ..., $s_n = i_n\sigma_n$, где коэффициенты $i_k = \pm 1$. ГСЗ более информативны, чем ГСЧ, поскольку их знаки несут дополнительную информацию о системе, в частности, разность числа положительных и отрицательных ГСЗ равна индексу Коши системы.

Далее будем рассматривать системы с кратными ГСЧ, причем ограничимся случаем, когда сингулярные числа принимают три различных значения.

Определение. *Система* (1), *ГСЧ которой принимают только три различных значения* σ_1, σ_2, σ_3, *называется трисингулярной. Ее передаточную функцию* $W(p)$ *также будем называть трисингулярной.*

В частности, любая система третьего порядка с различными ГСЧ является трисингулярной.

Блочно-сбалансированная декомпозиция ТДС

При наличии кратных сингулярных чисел σ_i у скалярной системы существует много сбалансированных представлений. Чтобы выделить среди них единственное, дополнительно потребуем диагональность кросс-грамиана системы. Полученная реализация будет представлять собой сбалансированную каноническую форму.

Опишем сбалансированную каноническую форму трисингулярных систем, следуя работам [3, 7]. Пусть $\sigma_1, \sigma_2, \sigma_3$ — ГСЧ системы, n_1, n_2, n_3 — их кратности, так что $n_1 + n_2 + n_3 = n$.

Матрицы $\mathbf{b}$ и $\mathbf{c}$ сбалансированной канонической формы имеют следующую структуру:

$$\mathbf{b}^T = \left[\underbrace{b_1,0,\ldots,0}_{n_1},\ \underbrace{b_2,0,\ldots,0}_{n_2},\ \underbrace{b_3,0,\ldots,0}_{n_3}\right];$$

$$\mathbf{c} = \left[\underbrace{i_1b_1,0,\ldots,0}_{n_1},\ \underbrace{i_2b_2,0,\ldots,0}_{n_2},\ \underbrace{i_3b_3,0,\ldots,0}_{n_3}\right],$$

где $i_k = \pm 1$, $k = 1, 2, 3$.

Таким образом, матрицы $\mathbf{b}$, $\mathbf{c}$ содержат по три блока

$$\mathbf{b} = \left[\mathbf{B}_1^T, \mathbf{B}_2^T, \mathbf{B}_3^T\right]^T,$$
$$\mathbf{c} = \left[\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3\right],$$
$$\mathbf{B}_k^T = \left[\mathbf{b}_k, 0, \ldots, 0\right],$$
$$\mathbf{C}_k = i_k\mathbf{B}_k^T,$$

размеры которых определяются кратностью сингулярных чисел.

Матрица $\mathbf{A}$ сбалансированной канонической формы также имеет блочную структуру, согласованную с блочной структурой матриц $\mathbf{b}$, $\mathbf{c}$. Ее диагональные блоки — квадратные матрицы $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ размерами n_1, n_2, n_3, они имеют трехдиагональный ленточный вид:

$$\mathbf{A}_k = \begin{bmatrix} a_{k1} & a_{k2} & & & & \\ -a_{k2} & 0 & a_{k3} & & & \\ & -a_{k3} & 0 & a_{k4} & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & & & 0 & a_{kn_k} \\ & & & & -a_{kn_k} & 0 \end{bmatrix}, \quad a_{k1} = -\frac{b_k^2}{2\sigma_k}. \tag{4}$$

Все диагональные элементы матрицы $\mathbf{A}_k$, кроме первого, равны нулю, элементы первой наддиагонали положительны, элементы первой поддиагонали и элемент a_{k1} — отрицательны. Известно, что подобную структуру имеют матрицы фазовращательных систем (матрицы Шварца).

Внедиагональные блоки $\mathbf{A}_{kj}$ матрицы $\mathbf{A}$ размеров $n_k \times n_j$ содержат по одному ненулевому элементу (он расположен в левом верхнем углу):

$$\mathbf{A}_{kj} = \begin{bmatrix} a_{kj} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} = i_k i_j \mathbf{A}_{kj}^T, \quad 1 \le k,\ j \le 3, \quad k \ne j,$$

причем элементы a_{kj} вычисляются по формуле

$$a_{kj} = \frac{-b_k b_j}{i_k i_j \sigma_k + \sigma_j}. \tag{5}$$

Таким образом, сбалансированная каноническая форма ТДС характеризуется следующими блочными матрицами:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_2 & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_3 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \mathbf{B}_3 \end{bmatrix}, \ \mathbf{c} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \mathbf{C}_3 \end{bmatrix}. \tag{6}$$

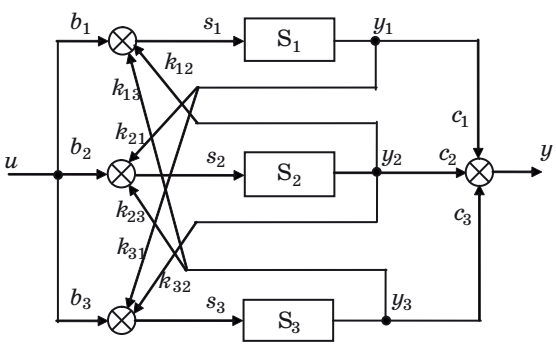
Им соответствует система уравнений в пространстве состояний:

$$\begin{aligned} \dot{\mathbf{X}}_1 &= \mathbf{A}_1\mathbf{X}_1 + \mathbf{A}_{12}\mathbf{X}_2 + \mathbf{A}_{13}\mathbf{X}_3 + \mathbf{B}_1 u, \ y_1 = \mathbf{C}_1\mathbf{X}_1; \\ \dot{\mathbf{X}}_2 &= \mathbf{A}_{21}\mathbf{X}_1 + \mathbf{A}_2\mathbf{X}_2 + \mathbf{A}_{23}\mathbf{X}_3 + \mathbf{B}_2 u, \ y_2 = \mathbf{C}_2\mathbf{X}_2; \\ \dot{\mathbf{X}}_3 &= \mathbf{A}_{31}\mathbf{X}_1 + \mathbf{A}_{32}\mathbf{X}_2 + \mathbf{A}_3\mathbf{X}_3 + \mathbf{B}_3 u, \ y_3 = \mathbf{C}_3\mathbf{X}_3, \\ y &= y_1 + y_2 + y_3. \end{aligned} \tag{6a}$$

Структурная схема, отвечающая этим уравнениям, показана на рис. 1. Она содержит три фазовращательные подсистемы S_1, S_2, S_3, соединенные обратными связями так, что выход каждого фазовращателя поступает на входы двух других.

Коэффициенты обратных связей $k_{qr} = \dfrac{-1}{i_q i_r \sigma_q + \sigma_r}$ на рис. 1 определяются значениями сингулярных чисел и не зависят от параметров самих фазовращателей.

Тем самым выполнена структурная декомпозиция ТДС на основе ее сбалансированного представления. Заметим, что внутренняя схемная реализация подсистем S_1, S_2, S_3 может быть любой



■ ***Рис. 1.*** *Блочно-сбалансированная декомпозиция ТДС*

и не обязательно должна описываться матрицами вида (4). Требуется лишь сохранить передаточные функции от входа u до выходов y_1, y_2, y_3 и систему сбалансированных обратных связей. Все это позволяет назвать полученную декомпозицию (см. рис. 1) блочно-сбалансированной.

По схеме рис. 1 можно построить передаточную функцию $W(p)$ ТДС, выразив ее через передаточные функции $\Phi_1(p)$, $\Phi_2(p)$, $\Phi_3(p)$ подсистем S_1, S_2, S_3 и числа s_1, s_2, s_3. Выписывая уравнения

$$y_1 = s_1\Phi_1(p)(-k_{12}y_2 - k_{13}y_3 + 1);$$
$$y_2 = s_2\Phi_2(p)(-k_{12}y_1 - k_{23}y_3 + 1);$$
$$y_3 = s_3\Phi_3(p)(-k_{13}y_1 - k_{23}y_2 + 1),$$
$$y = y_1 + y_2 + y_3$$

и исключая переменные y_1, y_2, y_3, получаем

$$W(p) = \frac{KQ_1Q_2Q_3 + 2(k_{12}Q_1Q_2 + k_{13}Q_1Q_3 + k_{23}Q_2Q_3) - (Q_1 + Q_2 + Q_3)}{2k_{12}k_{13}k_{23}Q_1Q_2Q_3 - (k_{12}^2Q_1Q_2 + k_{13}^2Q_1Q_3 + k_{23}^2Q_2Q_3) + 1}, \quad (7)$$

где использованы обозначения

$$Q_1 = s_1\Phi_1(p),\ Q_2 = s_2\Phi_2(p),\ Q_3 = s_3\Phi_3(p),\ \ K = k_{12}^2 + k_{13}^2 + k_{23}^2 - 2(k_{12}k_{13} + k_{12}k_{23} + k_{13}k_{23}).$$

Полученная блочно-сбалансированная форма позволяет решать задачу синтеза ТДС с заданными ГСЧ любой кратности. Соответствующий алгоритм был разработан и реализован в пакете MATLAB. Его исходными данными служат передаточные функции $\Phi_1(p)$, $\Phi_2(p)$, $\Phi_3(p)$ и значения трех ГСЗ s_1, s_2, s_3, результатом являются матрицы $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$ сбалансированного представления либо передаточная функция $W(p)$ ТДС. Кратность ГСЧ синтезированной системы определяется порядками используемых фазовращательных подсистем.

Циклические трисингулярные системы

Важный подкласс ТДС получаем в случае, когда все три базовые подсистемы в схеме рис. 1 одинаковы. Будем называть такие ТДС циклическими. Циклические трисингулярные системы (ЦТС) обладают рядом специальных свойств, обуславливаемых регулярностью и симметричностью их структуры.

Кратность всех трех сингулярных чисел ЦТС одинакова и равна порядку n_1 базовой подсистемы, а общий порядок системы равен $3n_1$. Размеры всех клеток матрицы $\mathbf{A}$ канонической формы (6) становятся одинаковыми, размеры клеток матриц $\mathbf{b}$, $\mathbf{c}$ также выравниваются.

Передаточная функция для циклических систем получается из формулы (7) при $\Phi_1(p) = \Phi_2(p) = \Phi_3(p) = \Phi(p)$:

$$W(p) = \frac{(s_1 + s_2 + s_3)\Phi\left(K\Phi^2 + 2\dfrac{k_{12}s_1s_2 + k_{13}s_1s_3 + k_{23}s_2s_3}{s_1 + s_2 + s_3}\Phi - 1\right)}{2k_{12}k_{13}k_{23}(s_1 + s_2 + s_3)\Phi^3 - (k_{12}^2s_1s_2 + k_{13}^2s_1s_3 + k_{23}^2s_2s_3)\Phi^2 + 1}. \quad (7a)$$

Циклические системы обладают важным свойством — независимостью формы диаграммы Найквиста от вида и порядка базовой фазовращательной подсистемы.

Простейшая ЦТС получится, если взять базовый фазовращатель первого порядка с передаточной функцией:

$$\Phi(p) = \frac{-p + a}{p + a} + 1 = \frac{2a}{p + a}.$$

Тогда матрицы канонической формы (6) принимают вид

$$\mathbf{A} = \begin{bmatrix} -a & a_{12} & a_{13} \\ a_{21} & -a & a_{23} \\ a_{31} & a_{32} & -a \end{bmatrix},\ \ \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},\ \ \mathbf{c} = \begin{bmatrix} i_1b_1 & i_2b_2 & i_3b_3 \end{bmatrix}.$$

Элементы матриц **A**, **b** и сингулярные числа связаны следующими соотношениями:

$$b_k = \sqrt{2a\sigma_k}, \quad a_{kj} = \frac{-2a\sqrt{\sigma_k\sigma_j}}{i_k i_j \sigma_k + \sigma_j}.$$

В частности, при $a = 1$ и положительных ГСЗ получаем

$$\mathbf{A} = -\begin{bmatrix} 1 & \frac{2\sqrt{\sigma_1\sigma_2}}{\sigma_1+\sigma_2} & \frac{2\sqrt{\sigma_1\sigma_3}}{\sigma_1+\sigma_3} \\ \frac{2\sqrt{\sigma_1\sigma_2}}{\sigma_1+\sigma_2} & 1 & \frac{2\sqrt{\sigma_2\sigma_3}}{\sigma_2+\sigma_3} \\ \frac{2\sqrt{\sigma_1\sigma_3}}{\sigma_1+\sigma_3} & \frac{2\sqrt{\sigma_2\sigma_3}}{\sigma_2+\sigma_3} & 1 \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} \sqrt{2\sigma_1} \\ \sqrt{2\sigma_2} \\ \sqrt{2\sigma_3} \end{bmatrix}, \quad \mathbf{c} = \mathbf{b}^T. \qquad (8)$$

Поскольку матрицы (8) полностью определяются ГСЗ, коэффициенты передаточной функции $W(p)$ такой системы также можно выразить через ГСЗ. Чтобы найти эту передаточную функцию, воспользуемся формулой

$$W(p) = \mathbf{c}(p\mathbf{E} - \mathbf{A})^{-1}\mathbf{b} = \frac{1}{\Delta}\mathbf{c}(p\mathbf{E} - \mathbf{A})^{*}\mathbf{b} = \frac{B(p)}{A(p)},$$

где **E** — единичная матрица; Δ — определитель матрицы $p\mathbf{E} - \mathbf{A}$; $B(p)$ и $A(p)$ — числитель и знаменатель искомой передаточной функции.

Выполняя вычисления, находим канонический вид передаточной функции циклической системы третьего порядка

$$W(p) = 2k\frac{p^2 + B_1 p + B_0}{p^3 + 3p^2 + A_1 p + B_0}, \qquad (9)$$

$$k = s_1 + s_2 + s_3;$$

$$B_1 = \frac{2}{s_1+s_2+s_3}\left(s_1\frac{s_1-s_2}{s_1+s_2} + s_2\frac{s_2-s_3}{s_2+s_3} + s_3\frac{s_3-s_1}{s_3+s_1}\right);$$

$$B_0 = \left(\frac{(s_1-s_2)(s_1-s_3)(s_2-s_3)}{(s_1+s_2)(s_1+s_3)(s_2+s_3)}\right)^2;$$

$$A_1 = \left(\frac{s_1-s_2}{s_1+s_2}\right)^2 + \left(\frac{s_1-s_3}{s_1+s_3}\right)^2 + \left(\frac{s_2-s_3}{s_2+s_3}\right)^2. \qquad (10)$$

К тому же результату приводит использование формулы (7а).

Рассмотрим задачу синтеза циклических систем третьего порядка с заданными ГСЗ s_1, s_2, s_3. Для ее решения можно использовать два подхода — матричный и структурный.

Первый подход опирается на формулы (8), которые позволяют получать матрицы описания в пространстве состояний циклической системы третьего порядка. Он дает матричное решение задачи синтеза для произвольного базового фазовращателя первого порядка. Формула (9) позволяет находить передаточную функцию полученной ЦТС.

Второй подход использует структурную схему рис. 1 и позволяет строить ЦТС при любом порядке базового фазовращателя.

Пример 1. Построим ЦТС третьего порядка с сингулярными числами $\sigma_1 = 2, \sigma_2 = 5, \sigma_3 = 9$.

Используя матричный подход и формулы (8), получаем

$$\mathbf{A} = \begin{bmatrix} -1 & -0{,}9035 & -0{,}7714 \\ -0{,}9035 & -1 & -0{,}9583 \\ -0{,}7714 & -0{,}9583 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 3{,}1623 \\ 4{,}2426 \end{bmatrix},$$

$$\mathbf{c} = \begin{bmatrix} 2 & 3{,}1623 & 4{,}2426 \end{bmatrix}.$$

Передаточная функция этой системы имеет вид

$$W(p) = 4\frac{47432p^2 + 20405p + 288}{5929p^3 + 17787p^2 + 3974p + 36}.$$

Подставляя коэффициенты этой передаточной функции в равенства (10), убеждаемся, что они выполняются, т. е. передаточная функция циклическая.

На рис. 2 приведена диаграмма Найквиста этой системы. Заметим, что ее горизонтальный размер равен удвоенной сумме сингулярных чисел, а площадь — сумме их квадратов, умноженной на π: $S = \pi(s_1^2 + s_2^2 + s_3^2) = 345{,}6$.

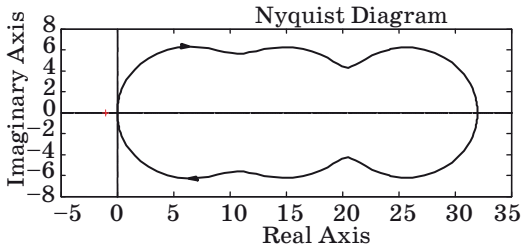
Пример 2. Синтезируем циклическую систему с ганкелевыми числами $s_1 = 1, s_2 = 2, s_3 = 3$.

Воспользуемся структурным подходом, взяв базовую подсистему с передаточной функцией $\Phi(p) = \frac{-p+2}{p+2} + 1 = \frac{4}{p+2}$. Вычисление коэффициентов обратных связей схемы рис. 1 дает

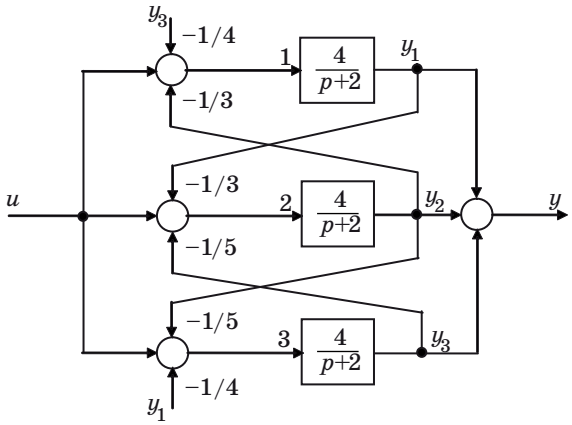
$$k_{12} = \frac{1}{s_1+s_2} = \frac{1}{3}, \quad k_{13} = \frac{1}{s_1+s_3} = \frac{1}{4},$$

$$k_{23} = \frac{1}{s_2+s_3} = \frac{1}{5}.$$

Это приводит к схеме блочно-сбалансированной декомпозиции, показанной на рис. 3.



■ ***Рис. 2.*** *Диаграмма Найквиста циклической системы с сингулярными числами 2, 5 и 9*



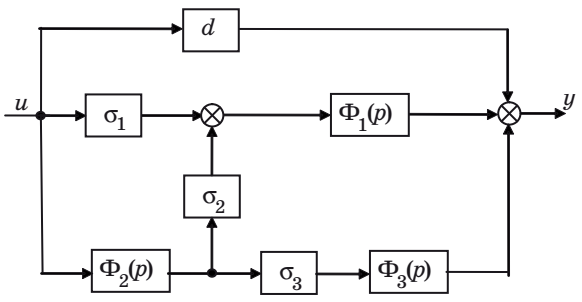
■ ***Рис. 3.*** *Структура циклической системы для примера 3*

Передаточную функцию этой системы получаем по формуле (7а):

$$W(p)=\frac{5400p^2+2760p+24}{225p^3+1350p^2+361p+2}.$$

Фазовая декомпозиция ТДС

Выше была получена блочно-сбалансированная декомпозиция ТДС (см. рис. 1). Ее вывод опирался на описание системы в пространстве состояний. Приведем другой вариант декомпозиции, который опирается на описание ТДС с помощью передаточной функции.



■ ***Рис. 4.*** *Фазовая декомпозиция ТДС*

В теории управления используют различные разложения передаточных функций, например разложение на простые дроби, разложение на устойчивую и антиустойчивую части и др. Менее известно фазовое разложение Гловера, в котором передаточная функция представляется в виде суммы фазовращательных слагаемых [2].

При этом в качестве коэффициентов при отдельных слагаемых выступают ГСЗ системы. Применение фазового разложения для декомпозиции бисингулярных систем описано в работе [7]. Рассмотрим особенности его применения для ТДС. В этом случае фазовое разложение описывается следующим образом.

Пусть $W(p)$ — устойчивая передаточная функция порядка n с ГСЧ $\sigma_1 > \sigma_2 > \sigma_3$ кратности r_1, r_2, r_3 соответственно, так что $r_1 + r_2 + r_3 = n$. Из теоремы Гловера [2] следует, что существует единственное представление $W(p)$ вида

$$W(p)=d+\sigma_1\Phi_1(p)+ \\ +\sigma_2\Phi_1(p)\Phi_2(p)+\sigma_3\Phi_2(p)\Phi_3(p), \qquad (11)$$

где d — константа; $\Phi_1(p)$, $\Phi_2(p)$, $\Phi_3(p)$ — устойчивые фазовращатели.

Отсюда вытекает, что любая устойчивая ТДС может быть представлена в виде соединения трех устойчивых фазовращателей согласно схеме, приведенной на рис. 4.

Передаточные функции фазовращателей $\Phi_1(p)$, $\Phi_1(p)$, $\Phi_3(p)$ имеют вид

$$\Phi_1(p)=\frac{A_1(-p)}{A_1(p)},\quad \Phi_2(p)=\frac{A_2(-p)}{A_2(p)},$$
$$\Phi_3(p)=\frac{A(-p)}{A(p)},$$

где $A_1(p)$, $A_2(p)$ — устойчивые полиномы; $A(p)$ — характеристический полином исходной системы.

Подставив эти передаточные функции в формулу (11), получаем

$$W(p)=d+\sigma_1\frac{A_1(-p)}{A_1(p)}+\sigma_2\frac{A_1(-p)}{A_1(p)}\cdot\frac{A_2(-p)}{A_2(p)}+ \\ +\sigma_3\frac{A_2(-p)}{A_2(p)}\cdot\frac{A(-p)}{A(p)}. \qquad (12)$$

Заметим, что соответствующая реализация неминимальна, поскольку сумма порядков всех фазовращателей превышает общий порядок системы. Тем не менее, разложение (12), в силу его единственности, можно считать канонической формой передаточной функции для устойчивых ТДС.

Для практического применения фазового разложения нужно уметь находить полиномы $A_1(p)$, $A_2(p)$ и константу d по заданной передаточной функции ТДС $W(p)$. При этом ГСЧ σ_1, σ_2, σ_3 в зависимости от постановки задачи могут задаваться заранее либо быть неизвестными. Порядки полиномов $A_1(p)$, $A_2(p)$ в общем случае считаются априорно неизвестными.

Процедура получения фазового разложения (12) достаточно трудоемка, особенно если часть параметров исходной передаточной функции задана в символьном виде. Решение можно получить методом неопределенных коэффициентов, приводя все члены уравнения (12) к общему знаменателю и приравнивая коэффициенты числителей при одинаковых степенях p справа и слева. Это приводит к системе нелинейных алгебраических уравнений относительно неизвестных параметров, решая которую, находим искомое разложение.

Для систем третьего порядка указанный алгоритм был реализован в пакете MAPLE и сводится к следующему.

Алгоритм фазового разложения передаточной функции $W(p)$ третьего порядка.

Шаг 1. Задаем элементы фазового разложения (12) в символьном виде.

Шаг 2. Переносим все члены разложения (12) в левую часть и приводим их, включая заданную передаточную функцию системы $W(p)$, к общему знаменателю.

Шаг 3. Выделяем в числителе коэффициенты при разных степенях p и приравниваем их нулю.

Шаг 4. Решаем полученную нелинейную систему уравнений, добавив ограничения для устойчивости полиномов $A_1(p)$, $A_2(p)$. Подставив полученное решение в фазовое разложение, заданное на шаге 1, получим ответ.

Наибольшую трудность при выполнении алгоритма представляет последний шаг, на котором необходимо решать нелинейную систему алгебраических уравнений с семью неизвестными.

Пример 3. Найдем фазовое разложение передаточной функции третьего порядка:

$$W(p)=\frac{12(900p^2+230p+1)}{900p^3+2700p^2+361p+1}.$$

Будем искать полиномы $A_1(p)$, $A_2(p)$ первого и второго порядков соответственно: $A_1(p) = p + a$, $A_2(p) = p^2 + c_1p + c_0$.

Для получения фазового разложения Гловера требуется найти значения семи параметров a, c_0, c_1, σ_1, σ_2, σ_3, d. Подставляем $W(p)$, $A_1(p)$, $A_2(p)$ в уравнение (12):

$$W(p)-\left(d+\sigma_1\frac{a-p}{a+p}+\sigma_2\frac{p^2-c_1p+c_0}{p^2+c_1p+c_0}\times\right.$$
$$\left.\times\frac{a-p}{a+p}+\sigma_3\frac{p^2-c_1p+c_0}{p^2+c_1p+c_0}\cdot\frac{A(-p)}{A(p)}\right)=0. \qquad (13)$$

Находим числитель левой части и выделяем коэффициенты при степенях p. Приравнивая их нулю и решая полученную систему нелинейных алгебраических уравнений, находим численные значения всех неизвестных:

$$a=\frac{1}{10},\ c_0=\frac{1}{150},\ c_1=\frac{5}{6},$$
$$\sigma_1=3,\ \sigma_2=2,\ \sigma_3=1,\ d=6.$$

Результат фазового разложения:

$$W(p)=6+3\frac{(1-10p)}{1+10p}+2\frac{(1-10p)(150p^2-125p+1)}{(1+10p)(150p^2+125p+1)}+$$
$$+\frac{(150p^2-125p+1)(-900p^3+2700p^2-361p+1)}{(150p^2+125p+1)(900p^3+2700p^2+361p+1)}.$$

Заметим, что константу d можно получить перед основными вычислениями. Подставим $p = 0$ и $p = \infty$ в формулу (12). Это дает два равенства

$$\frac{b_0}{a_0}=d+\sigma_1+\sigma_2+\sigma_3;\ \ \frac{b_n}{a_n}=d-\sigma_1-\sigma_2-\sigma_3.$$

Сложив их, получаем $d=\frac{1}{2}\left(\frac{b_n}{a_n}+\frac{b_0}{a_0}\right)$. В частности, в нашем примере имеем $d = 12/2 = 6$.

Синтез ТДС с заданным характеристическим полиномом

Фазовая декомпозиция ТДС (12) позволяет решать задачу синтеза ТДС, если в качестве исходных данных выступают характеристический полином системы и ее ГСЗ. В случае систем третьего порядка формальная постановка задачи сводится к следующему.

Даны ГСЗ s_1, s_2, s_3 и характеристический полином $A(p) = p^3 + a_2p^2 + a_1p + a_0$.

Требуется найти $B(p)$ — числитель искомой передаточной функции $W(p)=\frac{B(p)}{A(p)}$. Алгоритм решения, реализованный в пакете MAPLE, содержит 5 шагов.

Алгоритм синтеза ТДС третьего порядка с заданным характеристическим полиномом.

Шаг 1. Записываем фазовое разложение Гловера, включив d в состав $W(p)$.

В результате получаем выражение (13) при $d = 0$.

Шаг 2. Находим числитель правой части. Для этого приводим все слагаемые к общему знаменателю

$$\frac{s_1A_1(-p)A_2(p)A(p)+s_2A_1(-p)A_2(-p)A(p)}{A_1(p)A_2(p)A(p)}+$$
$$+\frac{s_3A_1(p)A_2(-p)A(-p)}{A_1(p)A_2(p)A(p)}=\frac{N(p)}{D(p)}.$$

Полином $N(p)$ должен делиться на $p + a$ нацело.

Кубический полином $A(p) = p^3 + a_2p^2 + a_1p + a_0$ и числа s_1, s_2, s_3 известны. Полином $N(p)$ шестого порядка должен делиться на произведение $A_1(p)$ $A_2(p)$ нацело, т. е. иметь вид $N(p) = A_1(p)A_2(p)B(p)$, где $B(p)$ — искомый полином третьего порядка. Выполняя деление $N(p)$ на $p + a$, находим частное Q и остаток R:

$$R = -2aA(a)[(a^2 + c)(s_1 + s_2) - ab(s_1 - s_2)].$$

Приравнивая нулю выражение в квадратных скобках, получаем уравнение связи неизвестных a, b, c, из которого выражаем c:

$$(a^2 + c)(s_1 + s_2) = ab(s_1 - s_2),$$
$$c = ab\frac{(s_1 - s_2)}{(s_1 + s_2)} - a^2. \qquad (14)$$

Шаг 3. Частное Q, полученное на шаге 2, должно делиться на полином p^2+bp+c нацело. Выполняя деление, получаем остаток в виде полинома от p первого порядка, его коэффициенты представляют собой полиномы от a, b третьего и четвертого порядков соответственно. Приравнивая их нулю, получаем два уравнения с двумя неизвестными: $P_1(a, b) = 0$, $P_2(a, b) = 0$.

Шаг 4. Находим результанты полиномов P_1, P_2 по a и b. После их факторизации и отбрасывания несущественных сомножителей получаем два полинома шестого порядка $R_1(a)$, $R_2(b)$. Вычисляя вещественные корни этих полиномов, находим искомые значения a, b.

Шаг 5. Для определения полинома $B(p)$ подставляем найденные значения a, b, c [последнее находим по формуле (14)] в частное Q, полученное на шаге 2. Этим завершается синтез передаточной функции $W(p)$ с заданным характеристическим полиномом и ГСЗ.

Замечание. Учет формулы (14) сокращает число неизвестных до двух и позволяет сразу начать с деления $N(p)$ на $A_2(p)$. Это уменьшает число шагов алгоритма.

Пример 4. Пусть требуется синтезировать систему третьего порядка с ГСЗ $s_1 = 3$, $s_2 = 2$, $s_3 = 1$ и характеристическим полиномом $A(p) = 900p^3 + 2700p^2 + 361p + 1$.

Шаг 1. Выписываем формулу фазового разложения Гловера

$$W(p)=3\frac{-p+a}{p+a}+2\frac{(-p+a)(p^2-bp+c)}{(p+a)(p^2+bp+c)}+$$
$$+\frac{(-900p^3+2700p^2-361p+1)(p^2-bp+c)}{(900p^3+2700p^2+361p+1)(p^2+bp+c)}$$

и находим полином $N(p)$ шестого порядка.

Шаг 2. Выполняя деление $N(p)$ на $p + a$, находим частное Q и остаток R:

$$R=-2a(361a-1-2700a^2+900a^3)(5a^2-ab+5c),$$

откуда $c=-a^2+\frac{1}{5}ab$.

Шаг 3. Выполняя деление частного Q на $p^2 + bp + c$, получаем остаток r — полином первого порядка. Его коэффициенты имеют вид

$$P_1=-13500b^4+(-12600a+13500)b^3+$$
$$+(51300a-41400a^2-5415)b^2+$$
$$+(7220a-36000a^3+5+67500a^2)b-$$
$$-25a(361a-1-2700a^2+900a^3);$$

$$P_2=4500a^3+(-13500+6300b)a^2+$$
$$+(3060b^2+1805-10800b)a-$$
$$-2700b^2-5+1083b+2700b^3.$$

Шаг 4. Находим результанты полиномов P_1, P_2 по a, b. Выполняя их факторизацию и отбрасывая несущественные сомножители, получаем

$$R_1(a)=(810000a^4-324000a^3+94221a^2-$$
$$-5220a+100)(-1+10a)^2;$$

$$R_2(b)=(29160000b^4-97200000b^3+$$
$$+118995300b^2-60813000b+12613381)\times$$
$$\times(-5+6b)^2.$$

Приравнивая результаты нулю, находим решение $a=\frac{1}{10}$, $b=\frac{5}{6}$, следовательно, $c=\frac{1}{150}$.

Шаг 5. Определение передаточной функции. Подставляя в частное Q найденные значения, получаем числитель $B(p)$ передаточной функции $W(p)$:

$$B(p)=-6p^3-6p^2+\frac{33}{50}p+\frac{1}{150}.$$

Таким образом, искомая передаточная функция имеет вид

$$W(p)=-6\frac{900p^3+900p^2-99p-1}{900p^3+2700p^2+361p+1}=$$

$$=-12\frac{900p^2+230p+1}{900p^3+2700p^2+361p+1}-6.$$

Заключение

В статье выделен и исследован класс линейных динамических систем, ГСЧ которых принимают ровно три значения. Получены две канонические декомпозиции таких систем — блочно-сбалансированная и фазовая. Соответствующие структурные схемы строятся на основе трех фазовращательных блоков. Отдельно рассмотрены циклические системы, характеризующиеся симметричной блочно-сбалансированной декомпозицией, и установлен вид их передаточной функции.

Предложены и реализованы алгоритмы синтеза ТДС с заданными ГСЧ, если дополнительно задан характеристический полином системы либо характеристические полиномы подсистем блочно-сбалансированной декомпозиции.

Полученные результаты могут быть полезны для решения задач идентификации, технической диагностики и редукции динамических систем.

Литература

1. **Francis B. A., Doyle J. C.** Linear control theory with on Hinf optimality criterion. A survey // SIAM J. Control and Optimization. 1987. Vol. 23. N 4. P. 815–844.
2. **Glover K.** All optimal Hankel-norm approximations of linear multivariable systems // Int. J. Control. 1984. Vol. 39. N 6. P. 1115–1193.
3. **Ober R. J.** Balanced parametrization of classes of linear systems // SIAM J. Control and Optimization. 1991. Vol. 29. N 6. P. 1251–1287.
4. **Мироновский Л. А.** Ганкелев оператор и ганкелевы функции линейных систем // Автоматика и телемеханика. 1992. № 9. С. 73–86.
5. **Мироновский Л. А.** Функциональное диагностирование динамических систем. — М.: Изд-во МГУ, 1998. — 340 с.
6. **Мироновский Л. А., Шинтяков Д. В.** Частотные характеристики фазовращательных и бисингулярных систем // Информационно-управляющие системы. 2007. № 5. С. 36–41.
7. **Мироновский Л. А.** Линейные системы с кратными сингулярными числами // Автоматика и телемеханика. 2009. № 1. С. 51–73.
8. **Пеллер В. В.** Операторы Ганкеля и их приложения: Пер. с англ. / НИЦ «РХД». — М.-Ижевск, 2005. — 1028 с.

УДК 621.391

КОДЫ ГОППЫ В ПРОТОКОЛАХ АНОНИМНОГО ЗАПРОСА К ДАННЫМ

С. В. Беззатеев,
канд. техн. наук, доцент
Санкт-Петербургский государственный университет аэрокосмического приборостроения

Предложен протокол анонимных запросов к данным, использующий семейства вложенных кодов Гоппы.

Ключевые слова *— распределенные базы данных, анонимность запросов, коды Гоппы.*

Введение

В настоящее время информация о предпочтениях и интересах частных лиц представляет определенный интерес и в некоторых случаях имеет вполне значимую цену. При этом, к сожалению, нельзя гарантировать, что владельцы информационных ресурсов будут соблюдать конфиденциальность и не будут собирать данные о предпочтениях клиентов с целью перепродать их третьим лицам. Протоколы извлечения информации без раскрытия запроса позволяют пользователю получить желаемую информацию из базы данных (БД) таким образом, что владелец БД ничего не узнает о номере бита, который запрашивал пользователь.

Понятие такого протокола впервые было введено в работе [1] под названием *Private Information Retrieval Protocol*, поэтому мы в дальнейшем будем называть такие протоколы «протоколами анонимного запроса к данным». Существует множество примеров, когда использование протоколов, которые скрывают от владельца БД интересы клиентов, пользующихся такими базами, может быть полезно. Приведем здесь лишь некоторые из таких систем.

Фармацевтические БД. Обычно фармацевтические компании интересуются либо вопросами изобретения и соответственно патентования новых лекарств, либо получением максимального объема информации об определенных компонентах лекарственных препаратов и их свойствах (фармацевтические БД). В процессе создания нового лекарства производителю требуется получить максимальный объем информации о свойствах его компонентов. По запросам сотрудников компании-производителя можно получить представление о том, разработкой какого типа лекарств занимается компания в данный момент. Чтобы скрыть планы компании, можно купить все необходимые для работы БД, однако это весьма дорого и, кроме того, требует постоянных обновлений. В этом случае использование протоколов анонимного запроса к данным позволяет избежать таких затрат.

Собственные распределенные БД. В настоящее время все более востребованными становятся распределенные хранилища данных, предоставляемые в пользование различным организациям. Такие услуги существенно снижают расходы компаний на хранение информации и поддержание работоспособности систем управления БД. Однако в таком случае появляется проблема конфиденциальности хранящейся и обрабатываемой таким образом информации. В этом случае защищаемой информацией является не только содержание хранимых и обрабатываемых данных, но и то, к каким данным обращаются пользователи наиболее часто или в какие моменты времени. Помимо этого, третьи лица могут интересоваться закономерностями в обращениях к определенным данным и определенным, наиболее частым связям запросов между собой. Использование для таких систем протоколов анонимного запроса позволяет обеспечить конфиденциальность не только самих данных в неконтролируемых хранилищах, но и скрыть информацию о запросах к таким БД.

Использование семейства вложенных кодов Гоппы для обеспечения анонимности запросов

Одним из способов построения протоколов анонимных запросов является использование помехоустойчивых кодов. В работах [2–4] для этих целей предлагается использовать специальный класс кодов — так называемые частично декодируемые

коды, т. е. коды, в которых можно восстанавливать определенную часть информационных символов. Здесь будет рассмотрен вариант построения протокола анонимных запросов, использующий семейства вложенных кодов Гоппы [5].

Данные, находящиеся в информационной базе, предварительно будут закодированы заранее выбранным множеством вложенных кодов Гоппы. Для каждого i-го поля записи выбирается свой код Гоппы с многочленом $G_i(x)$ и множеством нумераторов позиций L_i в качестве «базового». Информация из i-го поля кодируется с использованием надкода $(L_i, g_i(x))$, где $G_i(x) \equiv 0 \bmod g_i(x)$ и минимальные расстояния этих кодов соотносятся следующим образом:

$$d_i \le (d(\Gamma_i) - 1)/2,$$

где $d(\Gamma_i)$ — минимальное расстояние базового кода Гоппы $(L_i, G_i(x))$.

Таким образом, слова минимального веса надкода $(L_i, g_i(x))$ можно интерпретировать как векторы ошибок, которые могут быть исправлены «базовым $(L_i, G_i(x))$-кодом» Гоппы, выбранным для данного поля записи.

Принцип кодирования с использованием «базового» $(L_i, G_i(x))$-кода и его надкода $(L_i, g_i(x))$ выглядит в общем случае следующим образом. Информация, представленная кодовым словом $\mathbf{a}_1$ веса d_i надкода $(L_i, g_i(x))$, маскируется произвольным кодовым словом $\mathbf{b}_i$ «базового» кода Гоппы $(L_i, G_i(x))$:

$$\mathbf{e}_1 = \mathbf{a}_1 + \mathbf{b}_i.$$

В результате запись, состоящая из n различных полей, будет иметь вид $[\mathbf{e}_1\ \mathbf{e}_2\ \dots\ \mathbf{e}_n]$.

Очевидно, что в каждом поле информация представлена некоторым произвольным кодовым словом соответствующего надкода $(L_i, g_i(x))$. Запрос — *request* — на получение значения определенного k-го поля записи, обеспечивающий при этом анонимность запрашиваемого поля, будет выглядеть следующим образом:

$$request = [\mathbf{h}^*_1\ \mathbf{h}^*_2\ \dots\ \mathbf{H}^*_k\ \dots\ \mathbf{h}^*_n],$$

где $\mathbf{h}^*_i = \mathbf{A}_i \cdot \mathbf{h}_i \cdot \mathbf{P}_i$, здесь $\mathbf{A}_i$ — случайная матрица размером $r_G \times r_{gi}$, r_G — избыточность «базового» (L_i, G_i)-кода, r_{gi} — избыточность надкода $(L_i, g_i(x))$; $\mathbf{h}_i$ — проверочная матрица надкода размером $r_{gi} \times n_i$; $\mathbf{P}_i$ — перестановочная матрица размером $n_i \times n_i$, n_i — длина «базового» кода $(L_i, G_i(x))$;

$$\mathbf{H}^*_k = \mathbf{A}_k \cdot \mathbf{H}_k \cdot \mathbf{P}_k,$$

здесь $\mathbf{A}_k$ — случайная не особенная матрица размером $r_{Gk} \times r_{Gk}$; $\mathbf{H}_k$ — проверочная матрица «базового» кода размером $r_{Gk} \times n_k$; $\mathbf{P}_k$ — перестановочная матрица размером $n_k \times n_k$, n_k — длина надкода $(L_k, g_k(x))$, совпадающая с длиной «базового» $(L_k, G_k(x))$-кода.

Нетрудно доказать, что в результате выполнения запроса к информации, содержащейся в k-м поле записи при использовании выписанного вектора запроса *request*, получим

$$[\mathbf{e}_1\ \mathbf{e}_2\ \dots\ \mathbf{e}_n] \times [\mathbf{h}^*_1\ \mathbf{h}^*_2\ \dots\ \mathbf{H}^*_k\ \dots\ \mathbf{h}^*_n]^T = \mathbf{b}_k \cdot \mathbf{H}_k{}^{\mathrm{T}} \cdot \mathbf{A}_k{}^T,$$

где $\mathbf{b}_k$ — кодовое слово надкода $(L_k, g_k(x))$, вес которого не превышает корректирующей способности «базового» $(L_k, G_k(x))$-кода Гоппы.

Зная матрицу $\mathbf{A}_k$, легко получить значения синдромных компонент $\mathbf{R}_k$ для «базового» $(L_k, G_k(x))$-кода:

$$\mathbf{b}_k \cdot \mathbf{H}_k{}^T \cdot \mathbf{A}_k{}^T \cdot (\mathbf{A}_k{}^{-1})^T = \mathbf{b}_k \cdot \mathbf{H}_k{}^T = \mathbf{R}_k.$$

Далее, зная множество нумераторов позиций L_k и многочлен Гоппы $G_k(x)$ и применяя стандартный алгоритм декодирования для кодов Гоппы, можно вычислить вектор ошибок по его синдромным компонентам, т. е. фактически получить вектор $\mathbf{b}_k$, а следовательно, значение информационного вектора k-го поля. Таким образом мы восстановим значение запрашиваемого поля записи.

Заключение

Рассмотрен протокол, обеспечивающий анонимность запроса к полям БД, использующий свойства вложенных кодов Гоппы.

Литература

1. **Chor B., Goldreich O., Kushilevitz E., Sudan M.** Private information retrieval: Proc. of the 36th Annu. IEEE Symp. on Foundations of Computer Science. 1995. P. 41–51.
2. **Hemenway B., Ostrovsky R.** Public Key Encryption hich is Simultaneously a Loccaly-Decodable Error-Correcting Code // Electronic Collocquium on Computational Complexity. 2007. TR07-021. http://www.eccc.uni-trier.de (дата обращения: 03.11.2010).
3. **Yekhanin S.** New Locally Decodable Codes and Private Information Retrieval Schemes // Electronic Colloquium on Computational Complexity. 2006. TR06-127. http://eccc.hpi-web.de/eccc-reports/2006/TR06-127/index.html (дата обращения: 03.11.2010).
4. **Goldreich O., Karloff H., Schulman L., Trevisan L.** Lower bounds for linear locally decodable codes and private information retrieval systems: Proc. of the 17th IEEE Conf. on Complexity Theory. IEEE Computer Society Press, 2002. P. 263–296.
5. **Bezzateev S. V., Shekhunova N. A.** On the Subcodes of one class Goppa codes: Proc. Intern. Workshop Algebraic and Combinatorial Coding Theory (ACCT-1). Sept. 1988. P. 143–146.

XI МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ ПО ТЕЛЕКОММУНИКАЦИЯМ В ИНТЕЛЛЕКТУАЛЬНЫХ ТРАНСПОРТНЫХ СИСТЕМАХ — ITST-2011

www.itst2011.org

23–25 августа 2011 г.
Санкт-Петербург

Конференция ITST была учреждена в 2000 году Национальным институтом информационных и коммуникационных технологий (NICT, Япония) как один из первых специализированных форумов, посвященных тематике интеллектуальных транспортных систем (ИТС). За последние 10 лет аббревиатура ITST стала синонимом новейших идей и последних результатов исследований от ведущих представителей промышленности и науки, работающих в области информационно-коммуникационных технологий на транспорте во всем мире. Конференция проводилась в Японии (2000, 2001 и 2010), Корее (2002), Сингапуре (2004), Франции (2005, 2007 и 2009), Китае (2006), Таиланде (2008), а в 2011 году честь проведения этого форума предоставлена нашей стране.

Согласно официальной статистике ГИБДД, за 9 месяцев (январь—сентябрь) 2010 года в России произошло 143 608 дорожно-транспортных происшествий, в результате которых погибли 18 333 человека, а 181 779 человек получили ранения. Эти удручающие цифры указывают государственным властям всех уровней на необходимость решения таких жизненно важных проблем, как повышение эффективности контроля над соблюдением правил дорожного движения, улучшения качества дорожной инфраструктуры и уровня подготовки водителей. Все эти задачи в современном мире эффективно решаются с активным привлечением инфокоммуникационных технологий. Именно поэтому особое внимание на ITST в Санкт-Петербурге будет уделено тематике «Безопасность автомобильного транспорта».

Авторов приглашают принять участие в конференции с докладами по теории и практике автомобильных, железнодорожных, водных, авиационных и космических ИТС.

Ведущий организатор
Санкт-Петербургский институт информатики и автоматизации РАН (СПИИРАН)

Направления работы конференции
Законодательные, социальные и институциональные аспекты ИТС.
Коммуникационные архитектуры ИТС.
Приложения и сервисы ИТС на различных видах транспорта (автомобильный, включая общественный; водный; авиационный; железнодорожный).
Приложения кооперативных систем.
Спонтанные автомобильные сети (автомобиль-автомобиль, автомобиль — инфраструктура, географическая маршрутизация, широковещание и т. д.).
Мобильный IP и мобильность в IPv6.
Адресация и мобильность.
Протоколы безопасности в ИТС.
Управление рисками для сервисов безопасности (включая надежность систем).
Экологичные технологии ИТС.
Аналитическое и имитационное моделирование (NS2, NS3, OMNET, OPNET и т. д.).
Тестирование протоколов на соответствие требованиям, совместимость и оценка их качества.
Тестирование, верификация и диагностика компонентов ИТС.
Модели мобильности и автомобильного трафика.
Широковещательные мультимедиа технологии (TPEG, DVB, WiMAX, LTE и т. д.).
Электромагнитная совместимость.
Электрические автомобили и энергетическая инфраструктура.
Внутриавтомобильные сети.
Человекомашинные интерфейсы.
Автономное вождение.

Контрольные сроки
Подача статей на рецензирование — до 15 мая 2011 г.
Уведомление о принятии — до 30 июня 2011 г.
Подача окончательной версии для публикации — 15 июля 2011 г.

Дополнительная информация и справки
По всем вопросам (участие, спонсорство, организация выставок) обращайтесь: info@itst2011.org
Алексей Винель (СПИИРАН, Российская академия наук, Россия)
Официальный сайт мероприятия: www.itst2011.org
Мы будем рады приветствовать Вас в Санкт-Петербурге на ITST-2011!

IX МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ «ИДЕНТИФИКАЦИЯ СИСТЕМ И ЗАДАЧИ УПРАВЛЕНИЯ» — SICPRO'12

30 января – 2 февраля 2012 г.
Москва

Ведущий организатор
Институт проблем управления им. В. А. Трапезникова РАН

Цель конференции
Конференция SICPRO'12 имеет своей целью собрать вместе ученых, работающих во всех областях современной теории управления, для обсуждения таких вопросов, как: развитие теории и методологии идентификации, моделирования и управления; математические задачи теории управления; параметрическая идентификация; непараметрическая идентификация; структурная идентификация и экспертный анализ; задачи выбора и анализ данных; системы управления с идентификатором; задачи идентификации в интеллектуальных системах; прикладные задачи идентификации; имитационное моделирование; методическое и программное обеспечение идентификации и моделирования; когнитивные аспекты идентификации; верификация и проблемы качества программного обеспечения сложных систем; глобальные сетевые ресурсы поддержки процессов идентификации, управления и моделирования.

Условия участия
Конференция SICPRO'12 проводится:
— без регистрационного взноса;
— без ограничений на объем доклада.

Направления работы конференции
Структурная идентификация.
Параметрическая идентификация.
Непараметрическая идентификация.
Интеллектуальные методы идентификации.
Обработка сигналов.
Стохастические системы.
Адаптивные и робастные системы.
Методы оптимизации.
Приложения методов идентификации.

Контрольные сроки
Представление полных текстов докладов — не позднее 1 июля 2011 г.
Информация о принятии докладов — не позднее 1 октября 2011 г.
Объявление Программы конференции — не позднее 1 декабря 2011 г.
Заседания конференции SICPRO'12 — 30 января — 2 февраля 2012 г.

Дополнительная информация и справки
По всем организационным вопросам обращаться к Кириллу Чернышеву, Елене Ярко.
Эл. почта: noc@sicpro.org
Факс: + 7 (495) 334-89-90.
Телефон: + 7 (495) 334-89-90.
Официальный сайт конференции:
www.sicpro.org

СВЕДЕНИЯ ОБ АВТОРАХ

АНДРЕЕВ Сергей Дмитриевич

Научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН.
В 2006 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Комплексная защита объектов информатизации».
В 2009 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 40 научных публикаций.
Область научных интересов — беспроводные системы связи, системы массового обслуживания, мобильные и энергоэффективные системы.
Эл. адрес: corion@mail.ru

БЕЗЗАТЕЕВ Сергей Валентинович

Доцент кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1980 году окончил Ленинградский институт авиационного приборостроения по специальности «Автоматизированные системы управления».
В 1987 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 40 научных публикаций.
Область научных интересов - теория информации, теория кодирования, системы информационной безопасности.
Эл. адрес: bsv@aanet.ru

ВИНЕЛЬ Алексей Викторович

Старший научный сотрудник лаборатории информационных технологий в системном анализе и моделировании Санкт-Петербургского института информатики и автоматизации РАН.
В 2005 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по магистерскому направлению «Информационные системы в экономике».
В 2007 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 30 научных публикаций.
Область научных интересов — случайный множественный доступ, анализ и оценка производительности беспроводных сетей передачи данных.
Эл. адрес: vinel@ieee.org

АНИСИМОВ Алексей Валерьевич

Аспирант кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2007 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информационные системы и технологии».
Является автором девяти научных публикаций.
Область научных интересов — беспроводные сети передачи данных, обеспечение качества обслуживания, механизмы сбережения энергии.
Эл. адрес: alexey.anisimov86@gmail.com

БУДКОВ Виктор Юрьевич

Аспирант Санкт-Петербургского института информатики и автоматизации РАН.
В 2010 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Роботы и робототехнические системы».
Является автором восьми научных публикаций.
Область научных интересов — применение многомодальных интерфейсов в робототехнике и телекоммуникациях.
Эл. адрес: budkov@iias.spb.su

ГАЛИНИНА Ольга Сергеевна

Научный сотрудник научно-исследовательского департамента ООО «Центр речевых технологий», соискатель Санкт-Петербургского института информатики и автоматизации РАН.
В 2008 году окончила Санкт-Петербургский государственный политехнический университет по магистерскому направлению «Прикладная математика и информатика».
Является автором пяти научных публикаций.
Область научных интересов — математическая статистика, ЦОС, теория массового обслуживания и ее приложения, беспроводные технологии, биометрические технологии.
Эл. адрес: olga.galinina@gmail.com

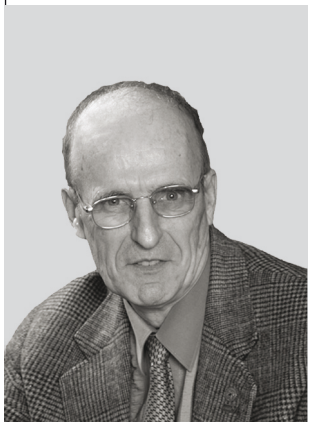
ЗЕЛЕНЦОВ
Вячеслав
Алексеевич

Ведущий научный сотрудник, руководитель группы космических информационных технологий и систем Санкт-Петербургского института информатики и автоматизации РАН, профессор Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1977 году окончил Военный инженерный Краснознаменный институт им. А. Ф. Можайского по специальности «АСУ и связь».
В 1993 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 180 научных публикаций и 21 изобретения.
Область научных интересов — системный анализ, теория надежности, методы управления эксплуатацией сложных систем.
Эл. адрес:
zvarambler@rambler.ru

КОТЕНКО
Игорь
Витальевич

Профессор, заведующий лабораторией проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН.
В 1983 году окончил Военный инженерный Краснознаменный институт им. А. Ф. Можайского, в 1987 году — Военную академию связи.
В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 450 научных публикаций, в том числе 12 учебников и монографий.
Область научных интересов — безопасность компьютерных сетей, в том числе анализ защищенности, обнаружение компьютерных атак, межсетевые экраны, защита от вирусов и сетевых червей, технологии моделирования и визуализации для противодействия кибер-терроризму.
Эл. адрес: ivkote@comsec.spb.ru

МИРОНОВСКИЙ
Леонид
Алексеевич



Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. Действительный член Академии навигации и управления движением, заслуженный работник высшей школы.
В 1962 году окончил Ленинградский политехнический институт.
В 1981 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 170 научных публикаций, соавтором пяти учебников и монографий, автором более 50 изобретений.
Область научных интересов — техническая диагностика и компьютерное моделирование динамических систем.
Эл. адрес: mir@aanet.ru

ЙОНДРАЛЬ
Фридрих

Гражданин ФРГ.
Профессор, директор Института связи (Institut für Nachrichtentechnik) Технологического института г. Карлсруэ, Германия.
В 1984 году защитил диссертацию на соискание ученой степени доктора наук.
Является автором более 200 научных статей.
Область научных интересов — UWB, программно-определяемые и интеллигентные радиосистемы, анализ сигналов, распознавание образов, оптимизация пропускной способности сетей и динамическое распределение каналов.
Эл. адрес:
friedrich.jondral@kit.edu

КУРМАЕВ
Илья
Ростомович

Аспирант кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 2006 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Информатика и вычислительная техника».
Является автором семи научных публикаций.
Область научных интересов — идентификация и компьютерное моделирование динамических систем.
Эл. адрес: leebowyer@mail.ru

НИКИТИН
Валерий
Николаевич

Доцент кафедры информационной безопасности телекоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. М. А. Бонч-Бруевича.
В 1982 году окончил Военную академию связи по специальности «Радиоэлектропроводная связь».
В 1991 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 50 научных и учебных публикаций и семи запатентованных изобретений.
Область научных интересов — системы радиосвязи и защиты информации, криптографические протоколы, методы согласования работы дискретных автоматов.
Эл. адрес: vnikitin@rdnet.ru

ОСИПОВ Дмитрий Сергеевич

Старший научный сотрудник лаборатории информационных технологий передачи, анализа и защиты данных Института проблем передачи информации им. А. А. Харкевича РАН.
В 2003 году окончил Московский государственный технический университет по специальности «Системы автоматического управления».
В 2008 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 10 научных публикаций.
Область научных интересов — теория передачи информации, разработка и исследование моделей систем множественного доступа, технологии защиты данных, передаваемых по беспроводным каналам связи, от подавления и прослушивания, методы оценивания характеристик беспроводных каналов связи.
Эл. адрес: d_osipov@iitp.ru

ПРИЩЕПА Мария Викторовна

Аспирантка Санкт-Петербургского института информатики и автоматизации РАН.
В 2010 году окончила Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Роботы и робототехнические системы».
Является автором четырех научных публикаций.
Область научных интересов — применение многомодальных интерфейсов в робототехнике.
Эл. адрес: prischepa@iias.spb.su

ТАНБУРГИ Ральф

Гражданин ФРГ.
Научный сотрудник института связи (Institut für Nachrichtentechnik) Технологического института г. Карлсруэ, Германия.
В 2010 году окончил Технологический институт г. Карлсруэ, Германия, по специальности «Информационная техника и электротехника».
Является автором семи научных публикаций.
Область научных интересов — одноранговые сети, маршрутизация на основе позиции и мобильной связи.
Эл. адрес: ralph.tanbourgi@kit.edu

ПАВЛОВ Александр Николаевич

Старший научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН, доцент Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1979 году окончил Ленинградский государственный университет им. А. А. Жданова по специальности «Математика».
В 1990 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 100 научных публикаций и трех изобретений.
Область научных интересов — системный анализ, теория принятия решений, теория надежности.
Эл. адрес: pavlov62@list.ru

РОНЖИН Андрей Леонидович

Доцент, заведующий лабораторией речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации РАН.
В 1999 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения.
В 2010 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 136 научных публикаций.
Область научных интересов — разработка речевых и многомодальных интерфейсов.
Эл. адрес: ronzhin@iias.spb.su

ТЮРЛИКОВ Андрей Михайлович

Доцент кафедры безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.
В 1980 году окончил Ленинградский институт авиационного приборостроения по специальности «Информационные системы управления».
В 1986 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 80 научных публикаций.
Область научных интересов — многоабонентные системы связи, системы дистанционного обучения, протоколы передачи данных в реальном масштабе времени, алгоритмы сжатия видеоинформации.
Эл. адрес: turlikov@mail.ru

ФРИДМАН Александр Яковлевич

Профессор, заведующий лабораторией информационных технологий управления промышленно-природными системами Института информатики и математического моделирования Кольского научного центра РАН.
В 1975 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина).
В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором 165 научных публикаций, в том числе одной монографии, 22 учебных пособий и 16 изобретений.
Область научных интересов — моделирование комплексных технологий и их воздействия на окружающую среду, прикладные интеллектуализированные системы.
Эл. адрес:
fridman@iimm.kolasc.net.ru

ЧЕЧУЛИН Андрей Алексеевич

Аспирант, младший научный сотрудник лаборатории проблем компьютерной безопасности Санкт-Петербургского института информатики и автоматизации РАН.
В 2005 году окончил Санкт-Петербургский государственный политехнический университет по магистерскому направлению «Информатика и вычислительная техника», специализация «Безопасность и защита информации».
Область научных интересов — безопасность компьютерных сетей, обнаружение вторжений, анализ сетевого трафика, анализ уязвимостей.
Эл. адрес:
chechulin@comsec.spb.ru

ЧУБИЧ Владимир Михайлович

Доцент кафедры прикладной математики Новосибирского государственного технического университета.
В 1984 году окончил Новосибирский электротехнический институт по специальности «Прикладная математика».
В 1996 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 40 научных публикаций, в том числе одной монографии.
Область научных интересов — анализ и планирование экспериментов для стохастических динамических систем.
Эл. адрес: chubich_62@ngs.ru

ФРИДМАН Ольга Владимировна

Старший научный сотрудник лаборатории информационных технологий управления промышленно-природными системами Института информатики и математического моделирования Кольского научного центра РАН.
В 1984 году окончила Петрозаводский государственный университет им. О. В. Куусинена по специальности «Физика».
В 1999 году защитила диссертацию на соискание ученой степени кандидата технических наук.
Является автором 95 научных публикаций, в том числе 13 учебных пособий.
Область научных интересов — прикладные интеллектуализированные системы, экспертные системы, нейронные сети.
Эл. адрес:
ofridman@iimm.kolasc.net.ru

ЧМОРА Андрей Львович

Ведущий специалист компании ОАО «Инфотекс».
В 1980 году окончил Московский электротехнический институт связи по специальности «Автоматическая электросвязь».
Является автором одной монографии и более 10 запатентованных изобретений.
Область научных интересов — криптография, теория информации, помехоустойчивое кодирование.
Эл. адрес:
andrey.chmora@mail.ru

ЭЛЬСНЕР Йенс

Гражданин ФРГ.
Научный сотрудник института связи (Institut für Nachrichtentechnik) Технологического института г. Карлсруэ, Германия.
В 2007 году окончил Технологический институт г. Карлсруэ, Германия, получив степень магистра.
Является автором 15 научных публикаций.
Область научных интересов — распределение каналов в одноранговых сетях, программно-определяемые и интеллигентные радиосистемы.
Эл. адрес: jens.elsner@kit.edu

Юркин Дмитрий Валерьевич

Аспирант кафедры информационной безопасности телекоммуникационных систем Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича.
В 2006 году окончил Санкт-Петербургский государственный университет телекоммуникаций им. М. А. Бонч-Бруевича по специальности «Защищенные системы связи».
Является автором более 20 научных и учебных публикаций.
Область научных интересов — системы радиосвязи и защиты информации, криптографические протоколы, методы согласования работы дискретных автоматов.
Эл. адрес: DVYurkin@yandex.ru

УВАЖАЕМЫЕ АВТОРЫ!

При подготовке рукописей статей редакция просит Вас руководствоваться следующими рекомендациями.

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 16 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала в Word шрифтом Times New Roman размером 13.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание, полное название организации, аннотация (7–10 строк) и ключевые слова на русском и английском языках, подрисуночные подписи.

Формулы в текстовой строке набирайте в Word, не используя формульный редактор (Mathtype или Equation), только в том случае, если средства Word не позволяют набрать формулу или символ (например, простая дробь, символы с «крышками» и т. д.), используйте имеющийся в Word формульный редактор Mathtype или Equation; формулы, стоящие в отдельной строке, могут быть набраны как угодно; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта никогда не пользуйтесь вкладкой Other..., используйте вкладку Define; в формулах не отделяйте пробелами знаки: + = –.

Для набора формул в Word никогда не используйте Конструктор (на верхней панели: «Работа с формулами» — «Конструктор»), т. к. этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Иллюстрации в текст не заверстываются и предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы изготавливаются в векторных программах: Visio 4, 5, 2002–2003 (*.vsd); Coreldraw (*.cdr); Excel; Word; AdobeIllustrator; AutoCad (*.dxf); Компас; Matlab (экспорт в формат *.ai);

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, факс, эл. адрес), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40 × 55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Более подробную информацию см. на сайте: www.i-us.ru

УДК 004.896

Система интеллектуального управления мобильным информационно-справочным роботом

Прищепа М. В., Будков В. Ю., Ронжин А. Л. Информационно-управляющие системы, 2010. № 6. С. 2–6.

Проанализирован круг проблем, возникающих при разработке обслуживающих информационных роботов. Предложена модель интеллектуального управления мобильной информационной системой на базе многомодального интерфейса, обеспечивающего естественное человекомашинное взаимодействие.

Ключевые слова — робототехника, системы интеллектуального управления, человекомашинное взаимодействие, мобильные подвижные системы.

Список лит.: 7 назв.

UDK 004.896

An intelligent control model for the mobile information-inquiry robot

Prischepa M. V., Budkov V. Yu., Ronzhin A. L. IUS, 2010. N 6. P. 2–6.

Problems of developing service information robots are discussed in this article. A model of intelligent control of mobile information system based on multimodal interface, providing natural human-robot interaction, is proposed.

Keywords — robotics, intelligent control systems, human-machine interaction, mobile moving systems.

Refs: 7 titles

УДК 519.8

Многокритериальный анализ влияния отдельных элементов на работоспособность сложной системы

Зеленцов В. А., Павлов А. Н. Информационно-управляющие системы, 2010. № 6. С. 7–12.

Предлагается метод решения задачи многокритериального анализа критичности отказов элементов сложной системы, основанный на комбинированном использовании метода нечеткого логического вывода и методов теории планирования эксперимента. Результирующий показатель критичности отказа элемента представляется в виде полинома, учитывающего влияние как отдельно взятых показателей, так и их совокупностей (по два, три и т. д.).

Ключевые слова — геном структуры, критичность отказа, сложный объект, многокритериальный анализ, теория планирования эксперимента, лингвистические переменные.

Список лит.: 15назв.

UDK 519.8

Multi-criteria analysis of separate elements influencing complex system performance

Zelentsov V. A, Pavlov A. N. IUS, 2010. N 6. P. 7–12.

A method of solving the task of multi-criteria failure criticality analysis of the complex systems elements' failures, based on a combined method of a fuzzy logic conclusion and of the experiment planning theory methods, is suggested. The resulting indicator of the element failure criticality is presented in the form of a polynom which accounts for both the influence of the separately taken indicators and the influence of the indicators' aggregations (of 2, 3, etc.).

Keywords — genome of structure, failure criticality, complex object, multi-criteria analysis, theory of experiment planning, linguistic variables.

Refs: 15 titles

УДК 004.9

Градиентный метод координации управлений иерархическими и сетевыми структурами

Фридман А. Я., Фридман О. В. Информационно-управляющие системы, 2010. № 6. С. 13–20.

Представлен градиентный метод координации децентрализованного управления иерархическими и сетевыми структурами на основе предложенных ранее необходимых и достаточных условий координируемости локально организованной иерархии динамических систем. Работоспособность метода проиллюстрирована результатами математического моделирования двухуровневой системы управления линейными объектами. Показано, что подключение локального управления и координации расширяет диапазон устойчивости системы к внешним и структурным возмущениям, а также повышает ее быстродействие в несколько раз.

Ключевые слова — ситуационный анализ и синтез, концептуальная модель предметной области, координируемость управляемых систем.

Список лит.: 9 назв.

UDK 004.9

A gradient coordination technique to control hierarchical and network systems

Fridman A. Ya., Fridman O. V. IUS, 2010. N 6. P. 13–20.

A gradient coordination technique for decentralized control in hierarchical and network systems is developed based on necessary and sufficient coordinability conditions, earlier proposed by authors for locally controlled hierarchies of dynamic systems. The efficiency of the technique is illustrated by mathematical modeling of a two-level control system for linear objects. The modeling has proved that usage of local control and coordination widens the stability range of the system for both external and structural disturbances as well as increases its speed several times.

Keywords — situational analysis and synthesis, conceptual model of subject domain, coordinability of controlled systems.

Refs: 9 titles

УДК 004.094

Комбинирование механизмов защиты от сканирования в компьютерных сетях

Чечулин А. А., Котенко И. В. Информационно-управляющие системы, 2010. № 6. С. 21–27.

Предлагается подход к комбинированию различных механизмов обнаружения сетевого сканирования в компьютерных сетях, который позволяет существенно повысить эффективность обнаружения за счет уменьшения количества ложных срабатываний и повышения точности обнаружения сканирования. Дается представление об отдельных частных методиках обнаружения сканирования. Рассматриваются основные принципы их комбинирования, а также дополнительные архитектурные улучшения общей модели обнаружения сканирования. Предлагается подход к автоматической настройке параметров используемых механизмов на основе статистических данных об анализируемом трафике.

Ключевые слова — защита информации, компьютерные сети, сетевое сканирование, обнаружение сетевых атак, комбинирование механизмов защиты.

Список лит.: 10 назв.

UDK 004.094

Combining scanning protection mechanisms in computer networks

Chechulin A. A., Kotenko I. V. IUS, 2010. N 6. P. 21–27.

An approach to combine different mechanisms of network scanning detection in computer networks is proposed. The approach allows to highly improve the detection effectiveness due to reduction of the false positive rate and increase in the scanning detection accuracy. Some particular scanning techniques are outlined. The core combining principles and architectural enhancements of the common scanning detection model are considered. An approach to automatically adjust the parameters of employed mechanisms based on statistical data about network traffic is also suggested.

Keywords — information security, computer networks, network scanning, network attack detection, combining of protection mechanisms.

Refs: 10 titles

УДК 621.391

Система множественного доступа, использующая некогерентный пороговый прием, частотно-позиционное кодирование и динамически выделяемый диапазон частот, в условиях подавления полезного сигнала

Осипов Д. С. Информационно-управляющие системы, 2010. № 6. С. 28–32.

Рассматривается модель системы множественного доступа, использующей динамически выделяемые поддиапазоны ортогональных частот, технологию частотно-позиционного кодирования и некогерентный пороговый прием в условиях подавления сигналов, передаваемых в системе. Для подавления используются сигналы, по форме и ширине занимаемой полосы аналогичные полезным. В работе проводится сравнение различных вариантов использования такой стратегии подавления с точки зрения их влияния на максимально возможную скорость надежной передачи информации рассматриваемым пользователем.

Ключевые слова — система множественного доступа, динамически выделяемые частотные поддиапазоны, некогерентный пороговый прием, подавление.

Список лит.: 4 назв.

UDK 621.391

On the performance of a non-coherent DHA FH OFDMA system with threshold reception under jamming

Osipov D. S. IUS, 2010. N 6. P. 28–32.

In this paper, the impact of intentional jamming on the performance of a Dynamic Hopset Allocation Frequency Hoping OFDMA system with non-coherent threshold reception is discussed. Jamming signals similar to those used by authorized users are considered. Various ways of implementing jamming strategy considered are compared in terms of maximum possible rate of the uplink channel corresponding to a certain authorized user.

Keywords — multiple access system, dynamically allocated hopsets, non-coherent threshold reception, intentional jamming

Refs: 4 titles

УДК 004.728.3.057.4

Оценка производительности простейшей системы абонентской кооперации

Андреев С. Д., Винель А. В., Галинина О. С. Информационно-управляющие системы, 2010. № 6. С. 33–41.

В работе рассмотрена модель системы абонентской кооперации, включающая трех абонентов. Получены замкнутые выражения для средней задержки передачи сообщения, а также для пропускной способности, энергетического потребления и энергетической эффективности абонентов-источников. Точность найденных характеристик подтверждается при помощи имитационного моделирования.

Ключевые слова — сотовая сеть, абонентская кооперация, система массового обслуживания, энергетическая эффективность.

Список лит.: 22 назв.

UDK 004.728.3.057.4

Performance evaluation of the simplest client relay network

Andreev S. D., Vinel A. V., Galinina O. S. IUS, 2010. N 6. P. 33–41.

This paper reviews a client relay system model comprising three nodes. Closed-form expressions for mean packet delay, as well as for throughput, energy consumption, and energy efficiency of the source nodes are obtained. The precision of the established parameters is verified by means of simulation.

Keywords — cellular network, client relay, queueing system, energy efficiency.

Refs: 22 titles

УДК 004.05

Улучшение способов аутентификации для каналов связи с ошибками

Никитин В. Н., Юркин Д. В. Информационно-управляющие системы, 2010. № 6. С. 42–46.

Рассмотрен обобщенный подход к разработке и анализу криптографических протоколов с помощью вероятностно-временных методов. Показано влияние ошибок, возникающих в канале связи, на работу протоколов аутентификации.

Ключевые слова — криптографический протокол, канал связи с ошибками, вероятностно-временные характеристики.

Список лит.: 14 назв.

UDK 004.05

Modification of authentification technics for error-prone channels

Nikitin V. N., Yurkin D. V. IUS, 2010. N 6. P. 42–46.

The article includes a description of the approach to the development and analysis of the cryptographic protocols using time-probability technique. It reveals the influence of channel errors on the performance of authentication and encryption protocols.

Keywords — cryptographic protocols, error-prone channel, time-probabilistic characteristics.

Refs: 14 titles

УДК 519.688

Кодовые шарады

Чмора А. Л. Информационно-управляющие системы, 2010. № 6. С. 47–53.

Рассматривается метод противодействия DoS-атаке с использованием шарад, построенных на основе кодов, исправляющих ошибки. Обосновывается практическая состоятельность итеративного метода конструирования шарад. Итеративные кодовые шарады позволяют исключить применение квантового компьютера и массированного распараллеливания в целях снижения трудоемкости отыскания решения.

Ключевые слова — DoS-атака, вычислительные шарады, коды, исправляющие ошибки, линейные коды.

Список лит.: 19 назв.

UDK 519.688

Code based puzzles

Chmora A. L. IUS, 2010. N 6. P. 47–53.

A countermeasure against connection depletion and other DoS attacks is discussed. We introduce a new client puzzle which is more robust to parallelization and quantum computer based attacks.

Keywords — denial of service, connection depletion attacks, client puzzles, error correcting codes, linear codes.

Refs: 19 titles

УДК 681.5.015

Активная параметрическая идентификация стохастических нелинейных непрерывно-дискретных систем на основе линеаризации во временной области

Чубич В. М. Информационно-управляющие системы, 2010. № 6. С. 54–61.

Впервые рассмотрены теоретические и прикладные аспекты активной параметрической идентификации стохастических нелинейных непрерывно-дискретных систем. Приведены оригинальные результаты для случая, когда подлежащие оцениванию параметры математических моделей могут входить в уравнения состояния и наблюдения, начальные условия и ковариационные матрицы помех динамики и ошибок измерений. Рассмотрен пример оптимального оценивания параметров одной модельной структуры.

Ключевые слова — оценивание параметров, метод максимального правдоподобия, планирование оптимальных входных сигналов, информационная матрица Фишера, критерий оптимальности.

Список лит.: 19 назв.

UDK 681.5.015

Active parametric identification of stochastic nonlinear continuous-discrete systems based on time domain linearization

Chubich V. M. IUS, 2010. N 6. P. 54–61.

Some theoretical and applied aspects of the active parametric identification of the stochastic nonlinear continuous-discrete systems are discussed for the first time. The original results are obtained for the case when the parameters of mathematical models to be estimated appear in the state and control equations, as well as the initial condition and the covariance matrices of the dynamic noise and measurement errors. An example of optimal parameter estimation for one model structure is shown.

Keywords — parameter estimation, maximum likelihood method, optimal input signal design, Fisher information matrix, optimality criterion.

Refs: 19 titles

УДК 004.728.3.057.4

Анализ влияния изменения характеристик потока на энергозатраты мобильной станции

Анисимов А. В., Тюрликов А. М. Информационно-управляющие системы, 2010. № 6. С. 62–69.

Рассматривается режим ожидания стандарта IEEE 802.16m, приводится анализ средней задержки и показателей энергоэффективности мобильной станции при приеме потока данных с переменной интенсивностью при использовании данного режима.

Ключевые слова — сбережение энергии, качество обслуживания, режим ожидания.

Список лит.: 18 назв.

UDK 004.728.3.057.4

An influence of data flow properties on energy consumption of mobile station

Anisimov A. V., Turlikov A. M. IUS, 2010. N 6. P. 62–69.

In the paper IEEE 802.16m sleep mode analysis is investigated. Mean delay and energy efficiency coefficients are analyzed for the case of data flow with varying rate receiving.

Keywords — energy saving, QoS, sleep mode.

Refs: 18 titles

УДК 004.07

О пропускной способности беспроводных многоканальных одноранговых сетей с местным планированием частотного разделения каналов

*Эльснер Й., Танбурги Р., Йондраль Ф.*Информационно-управляющие системы, 2010. № 6. С. 70–76.

Проведен анализ пропускной способности беспроводных многоканальных одноранговых сетей, ограниченных воздействием помех, с местным планированием частотного разделения каналов. Выведены верхняя и нижняя границы вероятностей прерывания связи и пропускной способности в данной модели системы. Выполнено сравнение с многоканальными сетями без местного планирования частотного разделения каналов.

Ключевые слова — многоканальные одноранговые сети, частотное разделение каналов, пропускная способность.

Список лит.: 10 назв.

UDK 004.07

About the transmission capacity of interference limited multi-channel ad hoc networks with local FDMA scheduling

Jondral F., Tanbourgi R., Elsner J. IUS, 2010. N 6. P. 70–76.

The transmission capacity of interference limited multi-channel ad hoc networks with local FDMA scheduling is analyzed. Upper and lower bounds on the success probability and the transmission capacity are derived and compared with the performance of multi-channel ad hoc networks without local FDMA scheduling.

Keywords — multi-channel ad hoc networks, FDMA, transmission capacity.

Refs: 10 titles

УДК 681.326.74

Синтез трисингулярных динамических систем

Мироновский Л. А., Курмаев И. Р. Информационно-управляющие системы, 2010. № 6. С. 77–85.

Выделен класс линейных стационарных систем с тремя группами кратных ганкелевых сингулярных чисел. Для таких систем, названных трисингулярными, установлен вид сбалансированного представления и найдены канонические структурные реализации. Разработан алгоритм декомпозиции трисингулярной передаточной функции на фазовращательные слагаемые и алгоритм синтеза систем с заданными ганкелевыми сингулярными числами и характеристическим полиномом.

Ключевые слова — линейные стационарные динамические системы, ганкелевы сингулярные числа, грамианы управляемости и наблюдаемости, декомпозиция передаточной функции, фазовращательные подсистемы.

Список лит.: 8 назв.

UDK 681.326.74

Synthesis of 3-singular dynamic systems

Mironovsky L.A., Kurmaev I. R. IUS, 2010. N 6. P. 77–85.

A class of linear time-invariant systems with three groups multiple Hankel singular values is alloted. For such systems, named 3-singular, a balanced representation is established, and canonical structural realizations is found. An algorithm of a decomposition of 3-singular transfer function on all-pass units is developed. An algorithm of synthesis of systems with given Hankel singular values and characteristic polynomial is designed as well.

Keywords — linear time-invariant systems, Hankel singular values, controllability and observability gramians, decomposition of transfer function, all-pass units.

Refs: 8 titles

УДК 621.391

Коды Гоппы в протоколах анонимного запроса к данным

Беззатеев С. В. Информационно-управляющие системы, 2010. № 6. С. 86–87.

Предложен протокол анонимных запросов к данным, использующий семейства вложенных кодов Гоппы.

Ключевые слова — распределенные базы данных, анонимность запросов, коды Гоппы.

Список лит.: 5 назв.

UDK 621.391

Goppa codes in anonymous data request protocols

Bezzateev S. V. IUS, 2010. N 6. P. 86–87.

A private information retrieval protocol, based on embedded Goppa codes, is described.

Keywords — distributed data bases, request anonymity, Goppa codes.

Refs: 5 titles

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (80x@mail.ru).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию. Рукописи не возвращаются.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

СОДЕРЖАНИЕ ЖУРНАЛА «ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ» ЗА 2010 г. [№ 1–6]

II МЕЖДУНАРОДНЫЙ КОНГРЕСС ПО СОВРЕМЕННЫМ ТЕЛЕКОММУНИКАЦИЯМ И СИСТЕМАМ УПРАВЛЕНИЯ – ICUMT-2010

18–20 октября 2010 г.
г. Москва, Россия

II Международный конгресс по современным телекоммуникациям и системам управления прошел 18–20 октября 2010 г. в Москве. Целью конгресса с момента его учреждения является интеграция молодых российских ученых в международное научное сообщество. В этот раз конгресс был организован Санкт-Петербургским институтом информатики и автоматизации РАН, Российским университетом дружбы народов и Технологическим университетом г. Тампере (Финляндия). Техническими спонсорами мероприятия выступили Институт инженеров по электротехнике и электронике (США) и Российское научно-техническое общество радиотехники, электроники и связи имени А. С. Попова.

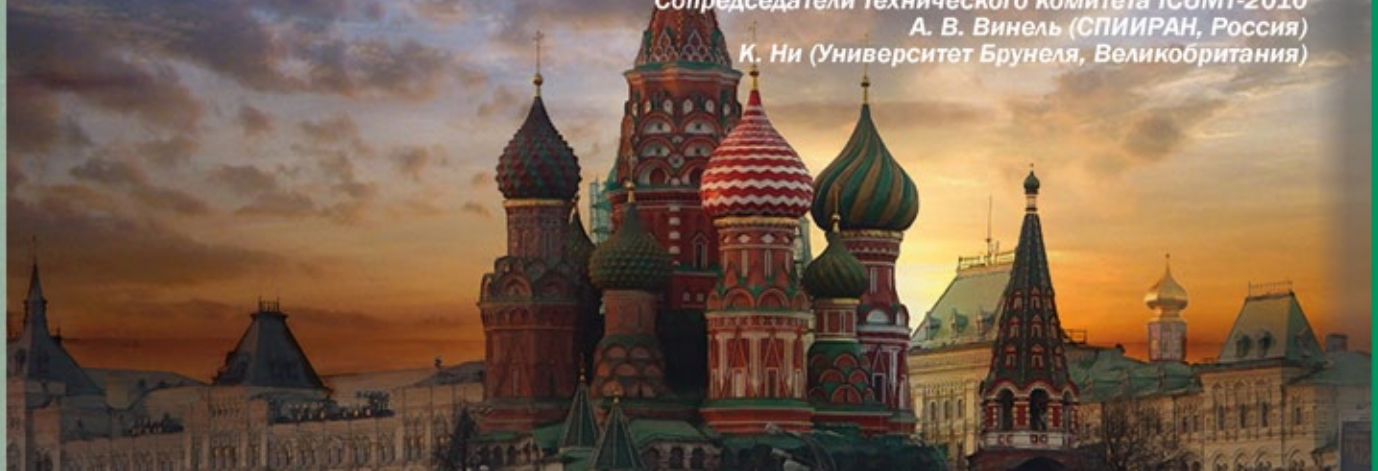
Техническая программа ICUMT-2010 включила в себя 5 ключевых докладов, 7 специализированных семинаров, промышленную секцию и 2 основных трека – "телекоммуникации" и "управление/робототехника". Членами технического комитета и рецензентами отобрано в труды конгресса свыше 200 публикаций из 320 поступивших. Конгресс носил поистине международный характер – в списке авторов 532 исследователя из 49 стран мира.

Ключевыми темами конгресса стали новейшие разработки в области беспроводной связи, а также результаты междисциплинарных исследований на стыке робототехники, систем управления и телекоммуникаций. Ключевые докладчики обрисовали состояние дел и представили прогнозы развития в области интеллектуальных транспортных систем, сетевой безопасности, многоканальной передачи, общества роботов и взаимодействия "человек–робот". Особенно большое количество публикаций относилось к тематике прикладных проблем: теории вероятности и математической статистики, мобильных вычислений, надежности сетей, оптических линий связи и, конечно, технологий беспроводной передачи данных. Промышленная секция "Будущее информационно-коммуникационных технологий: 2015 год" была организована приехавшими на конгресс руководителями научно-исследовательских подразделений компаний Hewlett-Packard (Франция), Alcatel-Lucent Bell Labs (Франция), Oracle (США), Nokia (Финляндия), Nokia Siemens Networks (Германия) и Telecom Italia (Италия).

Необходимо отметить тот факт, что если на I конгрессе ICUMT, проходившем в 2009 г. в Санкт-Петербурге, от нашей страны выступили 3–5% докладчиков, то в этом году доля их выросла до 20%: 40 из 200 статей, опубликованных в трудах, принадлежат ученым, живущим в России и представляющим отечественную науку. Труды конгресса доступны в Интернете в базе IEEE Xplore, а подробную информацию о конгрессе можно найти на сайте www.icumt.org.

Данный выпуск журнала включает в себя, в частности, расширенные и переведенные на русский язык избранные работы ICUMT-2010.

Сопредседатели технического комитета ICUMT-2010
А. В. Винель (СПИИРАН, Россия)
К. Ни (Университет Брунеля, Великобритания)



ВЫСТАВКА

АВТОМАТИЗАЦИЯ. ИНТЕЛЛЕКТУАЛЬНЫЕ ТЕХНОЛОГИИ. ИННОВАЦИИ

в рамках «VI Сибирского промышленного форума»

1–4 марта 2011 года

Автоматизация технологических процессов

Автоматизация производственных, коммерческих и жилых объектов

Инновационная обработка поверхностей

Логистика. Транспорт. Склад

VI Международная научно-практическая конференция

«ЛОГИСТИКА – ЕВРАЗИЙСКИЙ МОСТ»

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Сибирь
международный выставочно-деловой центр имени Карена Мурадяна

г. Красноярск, МВДЦ «Сибирь»,
ул. Авиаторов, 19
тел.: (391) 22-88-601
email: auto@krasfair.ru
www.krasfair.ru

ISSN 1684-8853

Официальная поддержка