БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

CSIST’2022

ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Материалы международного научного конгресса по информатике

В трех частях

Часть 1

Республика Беларусь
Минск, 27–28 октября 2022 г.

INFORMATION SYSTEMS AND TECHNOLOGIES

Proceedings of the International Scientific Congress on Computer Science

In three parts

Part 1

Republic of Belarus
Minsk, October 27–28, 2022

Научное электронное издание

МИНСК, БГУ, 2022

ISBN 978-985-881-424-3 (ч. 1)
ISBN 978-985-881-427-4

УДК 37:004(06)
ББК 74.044.4я431

Информационные системы и технологии = Information Systems and Technologies [Электронный ресурс] : материалы междунар. науч. конгресса по информатике. В 3 ч. Ч. 1, Респ. Беларусь, Минск, 27–28 окт. 2022 г. / Белорус. гос. ун-т ; редкол.: С. В. Абламейко (гл. ред.) [и др.]. – Минск : БГУ, 2022. – 1 электрон. опт. диск (CD-ROM). – ISBN 978-985-881-424-3.

Представлены материалы международного научного конгресса по информатике, организованного Белорусским государственным университетом и Объединенным институтом проблем информатики НАН Беларуси при поддержке ООО «Фабрика инноваций и решений».

Издание состоит из трех частей. В первой части рассматриваются следующие вопросы: информационная и компьютерная безопасность; интеллектуальный и статистический анализ данных, принятие решений; оптимизация и надежность систем.

Минимальные системные требования:
PC, Pentium 4 или выше; RAM 1 Гб; Windows XP/7/10;
Adobe Acrobat.

Оригинал-макет подготовлен в программе Microsoft Word.

На русском и английском языках

В авторской редакции

Ответственный за выпуск *И. С. Козловская*

Подписано к использованию 10.10.2022. Объем 4,3 МБ.

Белорусский государственный университет.
Управление редакционно-издательской работы.
Пр. Независимости, 4, 220030, Минск.
Телефон: (017) 259-70-70.
email: urir@bsu.by
http://elib.bsu.by

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ

Председатель конгресса

Абламейко Сергей Владимирович – академик Национальной академии наук (НАН) Беларуси, доктор технических наук

Сопредседатели

Медведев Дмитрий Георгиевич – профессор Белорусского государственного университета (БГУ), доктор педагогических наук, кандидат физико-математических наук

Тузиков Александр Васильевич – заведующий лабораторией Объединенного института проблем информатики НАН Беларуси (ОИПИ), член-корреспондент НАН Беларуси, доктор физико-математических наук

Харин Юрий Семенович – директор Научно-исследовательского института прикладных проблем математики и информатики БГУ, академик НАН Беларуси, доктор физико-математических наук

Председатель Программного комитета

Недзьведь Александр Михайлович – доктор технических наук, профессор, декан факультета прикладной математики и информатики БГУ

Председатель Организационного комитета

Казаченок Виктор Владимирович – заведующий кафедрой компьютерных технологий и систем, доктор педагогических наук, кандидат физико-математических наук

ПРОГРАММНЫЙ КОМИТЕТ

Председатель Программного комитета

Недзьведь Александр Михайлович – доктор технических наук, профессор, декан факультета прикладной математики и информатики БГУ, Беларусь

Члены программного комитета

Бодягин Игорь Александрович – кандидат физико-математических наук, доцент, заведующий кафедрой математического моделирования и анализа данных факультета прикладной математики и информатики БГУ, Беларусь

Босяков Сергей Михайлович – доктор физико-математических наук, профессор, декан механико-математического факультета БГУ, Беларусь

Бровка Наталья Владимировна – доктор педагогических наук, профессор, заведующий кафедрой теории функций БГУ, Беларусь

Васьковский Максим Михайлович – доктор физико-математических наук, доцент, заведующий кафедрой высшей математики факультета прикладной математики и информатики БГУ, Беларусь

Воротницкий Юрий Иосифович – кандидат физико-математических наук, доцент, заведующий кафедрой телекоммуникаций и информационных технологий факультета радиофизики и компьютерных технологий БГУ, Беларусь

Демиденко Олег Михайлович – доктор технических наук, профессор, проректор Гомельского государственного университета имени Ф. Скорины, Беларусь

Дмитрук Наталья Михайловна – кандидат физико-математических наук, доцент, заведующий кафедрой методов оптимального управления факультета прикладной математики и информатики БГУ, Беларусь

Журавков Михаил Анатольевич – доктор физико-математических наук, профессор, заведующий кафедрой теоретической и прикладной механики ММФ БГУ, Беларусь

Кадан Александр Михайлович – кандидат технических наук, доцент, заведующий кафедрой системного программирования и компьютерной безопасности Гродненского государственного университета имени Янки Купалы, Беларусь

Казаченок Виктор Владимирович – доктор педагогических наук, профессор, заведующий кафедрой компьютерных технологий и систем факультета прикладной математики и информатики БГУ, Беларусь

Каракозов Сергей Дмитриевич – доктор педагогических наук, профессор, проректор, директор Института математики и информатики Московского педагогического государственного университета, Россия

Ковалев Михаил Яковлевич – доктор физико-математических наук, профессор, член-корреспондент НАН Беларуси, заместитель генерального директора Объединенного института проблем информатики НАН Беларуси, Беларусь

Котов Владимир Михайлович – доктор физико-математических наук, профессор, заведующий кафедрой дискретной математики и алгоритмики БГУ, Беларусь

Краснопрошин Виктор Владимирович – доктор технических наук, профессор, заведующий кафедрой информационных систем управления факультета прикладной математики и информатики БГУ, Беларусь

Курбацкий Александр Николаевич – доктор технических наук, профессор, заведующий кафедрой технологий программирования факультета прикладной математики и информатики БГУ, Беларусь

Лиходед Николай Александрович – доктор физико-математических наук, профессор, профессор кафедры вычислительной математики факультета прикладной математики и информатики БГУ, Беларусь

Марков Сергей Викторович – кандидат физико-математических наук, доцент, заведующий кафедрой многопроцессорных систем и сетей факультета прикладной математики и информатики БГУ, Беларусь

Орлович Юрий Леонидович – кандидат физико-математических наук, доцент, заведующий кафедрой биомедицинской информатики факультета прикладной математики и информатики БГУ, Беларусь

Репников Василий Иванович – кандидат физико-математических наук, доцент, заведующий кафедрой вычислительной математики факультета прикладной математики и информатики БГУ, Беларусь

Русаков Александр Александрович – доктор педагогических наук, профессор, президент Академии информатизации образования, профессор МИРЭА-Российский технологический университет, Россия

Тузиков Александр Васильевич – доктор физико-математических наук, профессор, член-корреспондент НАН Беларуси, заведующий лабораторией Объединенного института проблем информатики НАН Беларуси, Беларусь

Харин Алексей Юрьевич – доктор физико-математических наук, профессор, заведующий кафедрой теории вероятностей и математической статистики факультета прикладной математики и информатики БГУ, Беларусь

Харин Юрий Семенович – доктор физико-математических наук, профессор, академик НАН Беларуси, директор Научно-исследовательского института прикладных проблем математики и информатики БГУ, Беларусь

Dao Van Tuyet – Professor, Binh Duong University, Vietnam

Demidenko Serge – Dean, Sunway University, Malaysia

Levashenko Vitaly – Head of Department, Zilina University, Slovakia

Marcelli Angelo – Professor, Salerno University, Italy

Seiichi Uchida – Professor, Fukuoka University, Japan

Ye Shiping – Vice-President, Zhejiang Shuren University, China

СОДЕРЖАНИЕ

УДК 51(09)

30 ЛЕТ РАЗВИТИЯ МАТЕМАТИКИ-ИНФОРМАТИКИ В РЕСПУБЛИКЕ БЕЛАРУСЬ: НЕКОТОРЫЕ РЕЗУЛЬТАТЫ

С.В. Абламейко

Объединенный институт проблем информатики НАН Беларуси, Минск, Беларусь;
Белорусский государственный университет, Минск, Беларусь

В работе рассматривается история развития математики-кибернетики-информатики в Республике Беларусь за период с 1992 по 2022 год. Приводятся основные факты, события, достижения.

Ключевые слова: математика; информатика; 30 лет; история развития.

30 YEARS OF DEVELOPMENT OF MATHEMATICS-INFORMATICS IN THE REPUBLIC OF BELARUS: SOME RESULTS

S.V. Ablameyko

United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus;
Belarusian State University, Minsk, Belarus

The paper analyses the history of the development of mathematics-cybernetics-informatics in the Republic of Belarus for the period from 1992 to 2022. The main facts, events, achievements are given.

Keywords: mathematics; informatics development; 30 years history.

Введение

В этом году исполнилось 30 лет с тех пор как мы живем в независимой Республике Беларусь. На это время пришелся период трудных 90-х годов, когда наука в Беларуси сильно сокращалась, перестраивалось и образование.

Советский союз и наша республика славились сильной математической наукой и математическим образованием. И нам, математикам, предстояло за эти 30 лет постараться не только не понизить эту планку, но и развивать уже белорусскую математику-информатику.

В статье кратко описывается пройденный путь, показываются наши результаты и достижения.

Советский период: очень кратко

В Советском союзе, особенно в последние десятилетия его существования, интенсивно развивалась кибернетика, как наука об общих закономерностях процессов управления и передачи информации в машинах, живых организмах и обществе. Это определение было предложено Норбертом Винером в 1948 году. Постепенно, где-то в 70-ые годы, стало все больше использоваться слово «информатика». «Энциклопедия Кибернетики», изданная в 1974 году в Киеве (тогда, когда кибернетика была в самом расцвете), определяла информатику как научную дисциплину, изучающую структуру и общие свойства научной информации. Сейчас информатика определяется, в основном, как наука об общих свойствах, закономерностях и методах обработки, поиска, передачи, хранения и использования информации с помощью компьютерных средств и технологий, а также область человеческой деятельности, связанная с их применением.

В советское время в БССР математику и кибернетику (информатику) преподавали и развивали в основном, в БГУ и МРТИ, а также в областных университетах (тогда они больше назывались педагогическими институтами). Особенно сильная школа по математике была в Гомельском государственном университете.

В АН БССР было два основных института, которые развивали математику и кибернетику. Институт математики был образован в 1959 году. Если во время создания Института математики в нем работало 4 доктора и 3 кандидата наук, то в 1980 году в Институте работали 504 человека, из них 202 научных сотрудника, включая 8 докторов и 45 кандидатов наук. В декабре 1990 г. из Института математики выделился в отдельную структуру Вычислительный центр АН БССР, который в 1995 году был присоединен к ИТК АН БССР.

В 1965 г. на базе лабораторий кибернетического профиля Института математики и вычислительной техники АН БССР образован Институт технической кибернетики. Во времена Советского Союза Институт был очень известной и весомой организацией. В 1984 году институт был награжден Орденом Трудового Красного Знамени. В 1991 году в институте работало более 1300 сотрудников.

Кроме этого, в БССР были очень сильные отраслевые институты и предприятия, такие как НИИ ЭВМ, НИИ СА – АГАТ, МПОВТ, МНИПИ, МНИРМ и другие.

Как известно, к концу советского периода БССР имела устойчивый имидж высокотехнологичной республики и у нас работала огромная армия ученых и практиков в сфере кибернетики и вычислительной техники.

Изменения за 30 лет

В Республике Беларусь все государственные организации, работающие в ИТ-секторе, можно разделить на несколько групп:

- университеты, где обучаются студенты и проводятся научные исследования в данной сфере;
- организации Национальной академии наук Беларуси, где проводятся научные исследования и выполняются разработки;
- отраслевые НИИ, ВЦ, выполняющие разработки для своей отрасли;
- а также все остальные организации нашей экономики, поскольку у большинства из них есть, как минимум, специалисты, обслуживающие вычислительную технику.

Все это было в начале 90-х годов прошлого века и работает сейчас также. Но понятно, что структура этих организаций и исследований значительно изменилась.

В 90-ые годы произошли большие сокращения. Наша маленькая республика (в сравнении с большим СССР) не могла содержать такую армию ученых. Часть людей уехала за рубеж, но это немного и в основном ученые. Большая часть людей ушла в бизнес. Немалая часть людей перешла в другие государственные структуры, поскольку у них стали повсеместно создаваться отделы вычислительной техники (назовем их так). Ну и немалая доля перешла в частный ИТ-бизнес.

Итак, что есть теперь.

Высшая школа

Исследования области математики-информатики проводятся во многих вузах Беларуси. Прежде всего – это БГУ, БГУИР, БНТУ, БГТУ и региональные классические университеты. За эти три десятилетия во всех региональных классических университетах расширились или открылись ИТ-факультеты (назовем их для упрощения так). Там работает достаточно большая армия ученых-исследователей, хотя докторов наук совсем немного.

В Беларуси стали открываться частные вузы. Первым частным вузом Беларуси был Институт современных знаний, основанный в октябре 1990 г. К середине 90-х гг. XX века число учреждений образования негосудар-

ственной формы собственности сильно увеличилось и их стало более 20-ти. В последние годы ряд частных вузов прекратил свое существование. Сейчас в Беларуси лишь 8 частных вузов.

К сожалению, в высшей школе наметилось старение кадров. Уменьшается количество преподавателей в возрасте до 39 лет и возрастает количество преподавателей старше 65 лет. Средний возраст докторов наук – примерно 69 лет.

В последние десятилетия стало уменьшаться количество кандидатов и докторов наук, работающих в высшей школе. Сокращается количество потенциальных научных руководителей аспирантов. Об этом ниже.

Национальная Академия наук Беларуси

В Академии наук две главные организации, где ведутся исследования по математике и информатике. Это Институт математики и Объединенный институт проблем информатики НАН Беларуси. Имеются и другие организации, но мы не будем их рассматривать.

Институт математики НАН Беларуси продолжает исследования на высоком уровне. В настоящее время в ИМ НАН Беларуси работает около 90 сотрудников, в том числе 16 докторов и 25 кандидатов наук.

Институт технической кибернетики в середине 1990-х годов был преобразован в ряд самостоятельных юридических структур, объединенных в НИО «Кибернетика». Сохранился собственно Институт в составе 330 человек, ведущий в основном теоретические исследования, и образовались четыре новых унитарных предприятия, ориентированные на создание научно-практических разработок: «Научное приборостроение», «Геоинформационные системы», «Информационные технологии», «Системы автоматизации».

В 2002 г. при реорганизации НИО «Кибернетика» и институт технической кибернетики был переименован в Объединенный институт проблем информатики, что больше соответствовало направлениям его деятельности. В его составе сейчас работают НИРУП «Геоинформационные системы» и НИРУП «Межотраслевой научно-практический центр систем идентификации и электронных деловых операций», созданное в 2006 г. Общая численность данных организаций составляет более 500 человек.

Численность сотрудников института (на 31.12.2021 г.) составляет 314 чел., в том числе 13 докторов наук и 55 кандидатов наук, научных сотрудников – 142.

ИТ-сектор

Особенно сильно изменился за эти 30 лет ИТ-сектор. В советский период существовали только государственные ИТ-организации, такие как НИИ ЭВМ, НИИ СА – АГАТ и другие.

Первые частные ИТ-компании появились в начале 90-х годов. Именно люди, работавшие до этого в сфере кибернетики и вычислительной техники в госструктурах стали создавать частные ИТ-предприятия, такие как IBA, EPAM-Systems и другие. И именно благодаря этим людям, ранее работавшим в госсекторе, появились и расцвели потом частные ИТ-компании.

В первое десятилетие своей деятельности белорусские компании работали больше по аутсорсинговой модели, по которой они попросту говоря разрабатывают какие-то кусочки, а весь продукт собирается за рубежом и продается от имени другой компании. Это было правильно и понятно, потому что только так можно набрать некоторую критическую массу для дальнейшего развития. За эти годы белорусские компании завоевали репутацию первоклассных разработчиков, действующих преимущественно по аутсорсинговой модели, но постепенно ориентиры меняются. Но где-то в конце нулевых годов стало понятно, что им надо переходить к продуктовой модели, когда весь продукт создается здесь и продается от имени белорусской компании.

Видя интенсивное развитие ИТ-бизнеса, государством в 2004 году был создан Парк высоких технологий, который успешно развивался. Новый импульс развитию ПВТ был дан в 2018 году после принятия Декрета Президента Республики Беларусь №8. Всего за пять лет с момента принятия цифрового Декрета количество резидентов Парка выросло в 5 раз и увеличилось со 192 до 1065. 2021 году экспорт ПВТ достиг абсолютного рекорда и составил $3,2 млрд [1].

Сейчас, по разным данным, в секторе информационно-коммуникационных технологий в Беларуси занято более 100 тысяч человек, из которых более 50 тысячи — в сегменте ИТ продуктов и услуг. Ещё около 50 тысяч ИТ-специалистов трудится в секторах экономики, отличных от ИКТ. В целом, отечественная инфраструктура ИКТ представляет собой стройную систему, существует хорошее государственно-частное партнерство. Парк высоких технологий, получивший определенный набор преференций, – яркий пример успешного развития информационно-телекоммуникационных технологий и взаимодействия науки, бизнеса и государства.

В ИТ-отрасли Беларуси отмечают большое количество сотрудников с высшим образованием — около 76 процентов. Другой характеристикой сектора является молодость — 57 процентов штата компаний-резидентов ПВТ имеют возраст до 30 лет. Карьерный путь в индустрии обычно начинается до 25 лет. Около 12 процентов занятых в ИТ-отрасли — студенты.

Можно сказать, что за прошедшие 15-20 лет практически создана новая отрасль. Но самое главное это то, что эта отрасль создана молодыми людьми, вчерашними выпускниками в основном двух университетов – БГУ и БГУИР. Средний возраст сотрудников, работающих в компаниях ПВТ – около 30 лет. Это очень молодые люди, совсем недавно закончившие университеты. И они не только работают простыми программистами под руководством старших товарищей. Многие компании возглавляются молодыми 30-летними людьми. Т.е. совсем еще молодые люди, по существу еще вчерашние выпускники, создали новую отрасль.

Это еще один аргумент о силе нашего высшего математического образования. Совершенно очевидно, что успехом развития ИТ-отрасли является хорошее образование с серьезной математической подготовкой. Это признавали и не раз отмечали оба директора Парка высоких технологий. Весь успех ИТ-компаний Беларуси базируется на хорошем образовании. А в основе этого образования, конечно, математика. И этим мы, математики, можем заслуженно гордиться.

Беларусь сейчас позиционируется в мире как страна с очень сильным ИТ-потенциалом.

Научные кадры

Надо сказать, что в девяностые годы белорусские ученые (те, кто остался в Беларуси) продолжали интенсивно работать и защищать докторские диссертации. Так продолжалось до начала этого века. Затем, в последнее десятилетие стало заметным старение и сокращение «кадров высшей квалификации». Это связано с нежеланием молодежи «идти и работать в науке», а также сложностью подготовки и защиты докторской диссертации.

Все эти прошедшие 30 лет, мне кажется, можно разбить на два периода по 15 лет. Как раз посредине – 2005-06 годы. До этого времени количество защит докторских диссертаций, по данным ВАК РБ, было примерно 100 в году. Потом начался спад. С 2006 года Высшая аттестационная комиссия (ВАК) Беларуси утверждает в среднем лишь 45–50 докторских диссертаций в год. В 2019 году было утверждено всего 39 докторских диссертаций. Столь низкого показателя по утвержденным степеням докторов наук в стране не было за три последних десятилетия.

Как отмечается в работе И. Шарыя [2] с 1995 по 2019 годы в Беларуси численность исследователей сократилась почти на четверть. За рассматриваемый период численность докторов наук устойчиво сокращалась и сократилась почти на 26 %. При этом, если в период с 2000 по 2010 гг. численность докторов наук сократилась на 9 %, то с 2010 по 2019 гг. – на 18,6 %. Таким образом, в последние годы темпы сокращения численности докторов существенно выросли.

Численность кандидатов наук за рассматриваемый временной период сократилась на 27,1 %. В результате сложившихся тенденций стала ухудшаться квалификационная структура исследователей. Если в 2000 г. доля НРВК составляла 23,7 % от общей численности исследователей, то в 2019 – 19,1 % [2].

По нашей информации, за последние 10 лет в Беларуси докторские диссертации в области фундаментальной и прикладной математики, информатики защитили только 5-7 человек.

В таблице 1 мы показываем, как изменилась численность докторов наук (математиков, информатиков) в основных вузах и организациях академии наук за прошедшие 30 лет. Данные являются примерными и взяты из открытых источников и опросов лидеров организаций и структур. Как видно из таблицы, в НАН Беларуси численность почти не изменилась, в то время как в вузах, численность докторов выросла.

Таблица 1. Количество докторов наук в организациях (без совместителей)

Организация	1992	2022
ФПМИ БГУ	9	18
ММФ БГУ	18	26
МРТИ-БГУИР	45	60
ИМ НАН Беларуси	17	16
ОИПИ НАН Беларуси	14	13
Итого	103	133

Вызывает большую тревогу тот факт, что доктора наук, профессора с каждым годом стареют, а молодых докторов наук прибавляется очень мало. В БГУ на двух математических факультетах средний возраст докторов наук составляет 68 лет. Ситуация не лучше и в других организациях. Докторов наук до 60 лет крайне мало. Понятно, что все это объясняется достаточно низким престижем ученых-преподавателей в стране, что привело молодых людей, особенно работающих в ИТ-сфере, к нежеланию связывать свою жизнь с наукой.

Показатели выпуска из аспирантуры и докторантуры с защитой диссертаций остаются на весьма низком уровне. Предпринимаемые вузами усилия по компенсации комплекса факторов, негативно влияющих на

системы подготовки и аттестации научных кадров высшей квалификации, не позволяют кардинально изменить ситуацию в системе послевузовского образования в силу ограниченности ресурсов (прежде всего материальных), имеющихся в распоряжении каждого вуза.

Признание заслуг внутри страны

За прошедшие 30 лет белорусские математики были удостоены многих наград Республики Беларусь, отмечены разными международными знаками отличия [3].

Высшая награда в области науки – Государственная премия Республики Беларусь была присуждена математикам за следующие работы:

1996 год – «Операторные методы в дифференциальных уравнениях» - Корзюк В.И., Радыно Я.В., Юрчук Н.И.

1998 год – «Модели и методы теории расписаний» - Танаев В.С., Сотсков Ю.Н., Гордон В.С., Струсевич В.А., Шафранский Я.М., Ковалев М.Я.

1998 год – «Учебное пособие «Лекции по теории графов» для высших и средних учебных заведений» - Емеличев В.А., Мельников О.И., Сарванов В.И., Тышкевич Р.И.

2000 год – «Исследование асимптотических свойств дифференциальных и дискретных систем» - Гайшун И.В., Изобов Н.А.

2002 год – «Распознавание и анализ стохастичеких данных и цифровых изображений» - Абламейко С.В., Харин Ю.С., Садыхов Р.Х., Старовойтов В.В., Тузиков А.В.

2004 год – «Метрическая теория диофантовых приближений зависимых величин и её приложения» - Берник В.И., Бересневич В.В.

Математики удостоены следующих наград Республики Беларусь:

Заслуженные деятели науки Республики Беларусь: С.В. Абламейко, Ф.М. Кириллова, М.М. Ковалев, А.Н. Курбацкий, Ю.С. Харин, Н.И. Юрчук.

Заслуженный работник образования Республики Беларусь: В.И. Корзюк, Р.И. Тышкевич. М.А. Журавков, В.В. Краснопрошин.

Кавалеры орденов и медалей Республики Беларусь: С.В. Абламейко, М.А. Журавков, В.В. Краснопрошин, В.И. Корзюк, П.А. Мандрик, Д.Г. Медведев, Н.И. Юрчук.

Членами Национальной академии наук Беларуси избраны следующие математики:

Академики НАН Беларуси: С.В. Абламейко, И.В. Гайшун, Н.А. Изобов, В.И. Корзюк, В.С. Танаев, С.А. Чижик, В.И. Янчевский. Ю.С. Харин.

Члены-корреспонденты НАН Беларуси: В.В. Гороховик, Ф.М. Кириллова, М.Я. Ковалев, Я.В. Радыно, А.В. Тузиков, Л.А. Янович.

Международное признание

Помимо наград и званий Республики Беларусь, ученым-математикам Беларуси были присуждены награды и звания других стран:

Анищенко В.В., Парамонов Н.Н. в 2006 году стали Лауреатами премии Правительства Российской Федерации в области науки и техники.

Л.А. Яновичу (совместно с украинскими коллегами) в 2012 году присуждена Государственная премия Украины.

С.В. Абламейко награжден в 2009 году Орденом Дружбы Российской Федерации.

И.В. Гайшун и В.Т. Борухов удостоены в 2012 году Премии Академий наук Украины, Беларуси и Молдовы.

С.В. Абламейко, В.В. Анищенко, С.В. Медведев удостоены в 2009 году премии Российской академии наук и Национальной академии наук Беларуси.

А.В. Тузикову, С.А. Золотому, С.А. Кореняко в 2021 году присуждена Межгосударственная премия «Звезды Содружества».

Кроме этого, белорусские математики избраны членам многих зарубежных академий, обществ и ассоциаций. Им присуждены звания почетных профессоров ряда зарубежных университетов. Они награждены различными медалями и орденами всевозможных ассоциаций и обществ. Мы не указываем эти награды и звания, поскольку очень трудно собрать их и особенно трудно понять их уровень и значимость. У многих профессоров это отмечено на их сайтах. Наиболее полная информация о наших математиках приведена на Портале «Выдающиеся математики Беларуси» [4].

Заключение

Подводя итог, можно сказать, что математика и информатика продолжала развиваться в новых условиях, сложившихся после распада Советского союза. Получено достаточно много новых результатов, в том числе, мирового уровня, защищено несколько десятков докторских диссертаций. Белорусские математики отмечены государственными наградами Республики Беларусь, а также других стран.

Все это сделано нынешним поколением ученых, которое, к сожалению, стареет. Мы гордимся тем, что, в этих непростых условиях, мы

смогли это сделать! И надеемся, что молодое поколение в какой-то степени сможет продолжить нашу работу.

Библиографические ссылки

1. ПВТ сегодня [Электронный ресурс]. URL: https://www.park.by/htp/about/.
2. Шарый И.Н. Особенности обеспечения стабилизации численности исследователей в научной сфере Республики Беларусь: социологический анализ // Социологический альманах / Институт социологии НАН Беларуси, РУП «Издательский дом «Беларуская навука». 2021. С. 156–166.
3. Абламейко С.В., Журавков М.А. Математика и математики БГУ и Беларуси. 100 лет развития. Минск: БГУ, 2021. 256 с.
4. Портал «Выдающиеся математики Беларуси» [Электронный ресурс]. URL: https://obm.bsu.by/.

ИНФОРМАЦИОННАЯ И КОМПЬЮТЕРНАЯ БЕЗОПАСНОСТЬ

УДК 512.713

АНАЛОГ КРИТЕРИЯ МИЛЛЕРА В ДЕДЕКИНДОВЫХ КОЛЬЦАХ С КОНЕЧНОЙ НОРМОЙ

М.М. Васьковский, Н.В. Кондратёнок

Белорусский государственный университет, Минск, Беларусь
vaskovskii@bsu.by, nkondr2006@rambler.ru

Найдены новые классы дедекиндовых колец с конечной нормой, в которых выполняется аналог критерия Миллера, а также получен эффективный алгоритм тестирования простоты идеалов, являющийся аналогом вероятностного теста Миллера-Рабина. В предположении справедливости расширенной гипотезы Римана доказан аналог теоремы Анкени, с помощью которой получено усиление аналога критерия Миллера, приводящее к детерминированному полиномиальному алгоритму тестирования простоты идеалов.

***Ключевые слова*:** дедекиндово кольцо; простые идеалы; критерий Миллера; тест на простоту.

AN ANALOGUE OF THE MILLER CRITERION IN DEDEKIND DOMAINS WITH A FINITE NORM PROPERTY

M.M. Vaskouski, N.V. Kondratyonok

Belarusian State University, Minsk, Belarus

There are found new classes of Dedekind domains with a finite norm property such that an analogue of the Miller criterion is valid in these domains, and an analogue of the Miller-Rabin algorithm for testing ideals primality is obtained. Assuming validity of the extended Riemann hypothesis, an analogue of Ankeny's theorem is proved that allows to obtain a strengthening of Miller's criterion analogue providing to a deterministic polynomial algorithm for testing the primality of ideals.

***Keywords*:** Dedekind domain; prime ideals; miller criterion; simplicity test; primality test.

Введение

Начиная со второй половины XX века начала активно развиваться информатика и криптография, что привело к увеличению активности работы в области алгоритмической теории чисел со стороны ведущих математиков. В частности, были получены принципиально новые критерии простоты, приводящие к эффективным алгоритмам тестирования простоты и генерации больших простых чисел. В 1980 году, опираясь на результат Г. Миллера [1], М.О. Рабиным [2] был получен безусловный полино-

миальный вероятностный алгоритм тестирования простоты, названный алгоритмом Миллера-Рабина. В предположении справедливости расширенной гипотезы Римана в работе [3] была доказана теорема, позволяющая получить детерминированный вариант алгоритма Миллера-Рабина. Алгоритм Миллера-Рабина играет ключевую роль при генерации ключей для RSA-криптосистемы.

Задача проверки на простоту существенно усложняется при переходе от целых чисел к более общим алгебраическим структурам. В работе [4] был получен полиномиальный детерминированный алгоритм проверки на простоту в конечнопорожденных дедекиндовых кольцах. В работе М.М. Васьковского, Н.В. Кондратёнка и Н.П. Прохорова [5] был разработан подход, позволяющий переносить доказательства ряда известных критериев простоты на различные алгебраические структуры. В 2020 году в работе Н.П. Прохорова [6] были доказаны новые критерии простоты в кольцах целых алгебраических чисел. В настоящей статье доказывается аналога критерия Миллера в некотором классе дедекиндовых колец, а также приводится соответствующий алгоритм тестирования простоты идеалов.

1. Основные результаты

Пусть R дедекиндово кольцо. Нормой $Nm(\mathfrak{n})$ идеала $\mathfrak{n} \subset R$ называется мощность факторкольца $R/\mathfrak{n}$. Говорят, что дедекиндово кольцо R является дедекиндовым кольцом с конечной нормой (finite norm property), если для любого собственного идеала $\mathfrak{n} \subset R$ факторкольцо $R/\mathfrak{n}$ конечно. Далее в работе будем рассматривать только дедекиндовы кольца с конечной нормой. Пусть $a, b \in R$ и $\mathfrak{n} \subseteq R$. Будем говорить, что a сравнимо с b по модулю $\mathfrak{n}$ и писать $a \equiv b(mod\ \mathfrak{n})$, если $a - b \in \mathfrak{n}$.

Определение 1. [7, с. 285] Функцией Эйлера нетривиального идеала $\mathfrak{n} \subset R$ называется функция

$$\varphi(\mathfrak{n}) = \left|I_{R/\mathfrak{n}}\right|.$$

Определение 2. Элемент $a \in R$ будем называть квадратичным вычетом по модулю идеала $\mathfrak{n}$, если существует $b \in R$, что $b^2 \equiv a(mod\ \mathfrak{n})$.

Для простого идеала $\mathfrak{p}$ и $a \in I_{R/\mathfrak{p}}$ определим символ Лежандра следующим образом

$$\left(\frac{a}{\mathfrak{p}}\right) = \begin{cases} 1, \text{если } a \text{ квадратичный вычет по модулю } \mathfrak{p}, \\ -1, \text{иначе.} \end{cases}$$

Для нетривиального идеала $\mathfrak{n} = \mathfrak{p}_1 \dots \mathfrak{p}_k$ и $a \in I_{R/\mathfrak{n}}$ определим символ Якоби следующим образом

$$\left(\frac{a}{\mathfrak{n}}\right) = \left(\frac{a}{\mathfrak{p}_1}\right) \dots \left(\frac{a}{\mathfrak{p}_k}\right).$$

Пусть K является полем частных R. Пусть расширение $L \supset K$ конечное, сепарабельное и нормальное расширение, а $Gal(L/K)$ является абелевой. Положим S алгебраическое замыкание R в L.

Определение 3. Пусть $\mathfrak{p}$ простой идеал кольца R. Рассмотрим идеал $\mathfrak{p}S$, который он генерирует в кольце S. Рассмотрим его факторизацию на простые идеалы

$$\mathfrak{p}S = \prod\nolimits_{\mathfrak{q}} \mathfrak{q}^{e_{\mathfrak{q}}},$$

где произведение берется по всем простым идеалам кольца S и $e_{\mathfrak{q}} > 0$ только для конечного количества простых $\mathfrak{q}$.

Если $e_{\mathfrak{q}} > 0$ для некоторого $\mathfrak{q}$, то говорят, что $\mathfrak{q}$ лежит над (lie over, lie above) простым идеалом $\mathfrak{p}$. Число $e_{\mathfrak{q}}$ из разложения называется индексом разветвления (ramification index) $\mathfrak{q}$.

Если для простого идеала $\mathfrak{p} \subseteq R$ выполнено $e_{\mathfrak{q}} > 1$ для некоторого $\mathfrak{q} \subseteq R$, то говорят, что идеал $\mathfrak{p}$ ветвится в L. Если идеал $\mathfrak{p}S$ простой в S, то говорят, что $\mathfrak{p}$ инертный (inert). Если для всех $\mathfrak{q}$ выполняется или $e_{\mathfrak{q}} = 0$, или $e_{\mathfrak{q}} = 1$, то говорят, что $\mathfrak{p}$ полностью разлагается (splits completely) в L.

Определение 4. Пусть $\mathfrak{p}$ простой идеал кольца R, не ветвящийся в L и пусть $\mathfrak{P} = \mathfrak{p}S$ соответствующий идеал в S. Тогда существует единственный такой элемент $\sigma \in Gal(L/K)$, что для любого $\alpha \in L$ выполнено $\sigma(\alpha) \equiv \alpha^{Nm(\mathfrak{p})} (mod\ \mathfrak{P})$. Этот элемент называют символом Артина идеала $\mathfrak{p}$.

Определение 5. Пусть $\phi: Gal(L/K) \to I_{\mathbb{C}}$ гомоморфизм. Рассмотрим функцию

$$\chi(\mathfrak{p}) = \begin{cases} \phi\left(\sigma_{\mathfrak{p}}\right), \text{если } \mathfrak{p} \text{ не ветвится}, \\ 0, \text{иначе}, \end{cases}$$

где $\mathfrak{p}$ простой и $\sigma_{\mathfrak{p}}$ символ Артина идеала $\mathfrak{p}$. Используя мультипликативность, эту функцию можно определить для всех идеалов R. Полученную функцию χ будем называть характером. Характер, принимающий только значения 0 и 1, называется главным.

Определение 6. Будем говорить, что характер χ задан по модулю идеала $\mathfrak{f} \subset R$, если для всех идеалов $\mathfrak{n} \subseteq R$, из сравнения $\mathfrak{n} \equiv 1 (mod\ \mathfrak{f})$ следует равенство $\chi(\mathfrak{n}) = 1$.

Определение 7. Пусть характер χ задан на множестве идеалов кольца R, не является главным и определен по модулю идеала $\mathfrak{n} \subset R$. Через $\mathfrak{p}_\chi$ обозначим идеал минимальной нормы, для которого $\chi\left(\mathfrak{p}_\chi\right) \neq 0, 1$.

Определение 8. Пусть R дедекиндово кольцо с полем частных K и L расширение поля K степени не меньше 2. Будем говорить, что кольцо R удовлетворяет условию A для идеала $\mathfrak{n}$, если существует многочлен f_R, что для любого характера χ, не являющегося главным и определенного по модулю $\mathfrak{n}$, выполнено

$$Nm(\mathfrak{p}_\chi) \leq f_R(\log Nm(\mathfrak{n})).$$

Замечание. Из работы [8] следует, что, если расширенная гипотеза Римана выполнена, то условие A выполнено для всех колец $\mathcal{O}_K$ целых алгебраических чисел числового поля K и $f_{\mathcal{O}_K}(x) = 12\mathrm{x}^2 + 12\log^2 \Delta_K$.

Замечание. Из работы [3] следует, что, если обобщенная гипотеза Римана выполнена, то условие A выполнено для кольца целых чисел и $f_{\mathbb{Z}}(x) = 2x^2$.

Предложение 1. Пусть кольцо R удовлетворяет условию A. Пусть $\chi: I_{R/\mathfrak{n}} \to G$ нетривиальный гомоморфизм. Тогда существует идеал $\mathfrak{a}$ взаимнопростой с $\mathfrak{n}$ и такой, что $\chi(\mathfrak{a}) \neq 1$ и $Nm(\mathfrak{a}) \leq f_R(\log Nm(\mathfrak{n}))$.

Доказательство. Из условия предложения следует, что подгруппа $\chi(I_{R/\mathfrak{n}}) \subseteq G$ нетривиальная. Рассмотрим нетривиальный характер $\xi: \chi(I_{R/\mathfrak{n}}) \to I_{\mathbb{C}}$. Очевидно, что $\xi \circ \chi: I_{R/\mathfrak{n}} \to I_{\mathbb{C}}$ является нетривиальным характером группы $I_{R/\mathfrak{n}}$.

Из определения условия A следует, что существует простой $\mathfrak{a}$ взаимнопростой с $\mathfrak{n}$ и такой, что $(\xi \circ \chi)(\mathfrak{a}) \neq 1$ и $Nm(\mathfrak{a}) \leq f_R\big(\log\big(Nm(\mathfrak{n})\big)\big)$. Из того, что $(\xi \circ \chi)(\mathfrak{a}) \neq 1$ следует, что $\chi(\mathfrak{a}) \neq 1$.

Предложение 2. Пусть идеал $\mathfrak{p}$ простой с нечетной нормой. Тогда сравнение $x^{Nm(\mathfrak{p})-1} \equiv 1 (mod\ \mathfrak{p}^2)$ имеет не более $Nm(\mathfrak{p}) - 1$ решений относительно $x \in I_{R/\mathfrak{p}^2}$.

Доказательство. Из теоремы Эйлера [7, с. 285] следует, что сравнение $x^{Nm(\mathfrak{p})-1} \equiv 1 (mod\ \mathfrak{p})$ имеет ровно $Nm(\mathfrak{p}) - 1$ решений относительно $x \in I_{R/\mathfrak{p}}$.

Заметим, что все решения сравнения $x^{Nm(\mathfrak{p})-1} \equiv 1 (mod\ \mathfrak{p}^2)$ имеют вид $a + \mathfrak{p}\mathfrak{t}$, где $x \in I_{R/\mathfrak{p}}$, $\mathfrak{t} \in R/\mathfrak{p}$ и a является решением сравнения $x^{Nm(\mathfrak{p})-1} \equiv 1 (mod\ \mathfrak{p})$. Подставим этот вид в сравнение, раскроем скобки и получим сравнение $\mathfrak{p}\mathfrak{t}(Nm(\mathfrak{p}) - 1)a^{Nm(\mathfrak{p})-2} \equiv 1 - a^{Nm(\mathfrak{p})-1} (mod\ \mathfrak{p}^2)$. Так как $\Big((Nm(\mathfrak{p}) - 1)a^{Nm(\mathfrak{p})-2}, \mathfrak{p}\Big) = 1$, то это сравнение имеет ровно одно решение при фиксированном a. Следовательно, исходное сравнение имеет не более $Nm(\mathfrak{p}) - 1$ решений относительно $x \in I_{R/\mathfrak{p}^2}$.

Из рассуждений работы [6] вытекает следующий аналог критерия Эйлера для дедекиндовых колец с конечной нормой.

Утверждение 2. Пусть $\mathfrak{n}$ – нетривиальный идеал нечетной нормы дедекиндового кольца R. Тогда $\mathfrak{n}$ – простой идеал тогда и только тогда, когда для любого $a \in I_{R/\mathfrak{n}}$ выполнено $a^{\frac{Nm(\mathfrak{n})-1}{2}} \equiv \left(\frac{a}{\mathfrak{n}}\right) (mod\ \mathfrak{n})$.

Теорема 1. Пусть $\mathfrak{n}$ – нетривиальный идеал нечетной нормы дедекиндового кольца R. Пусть $Nm(\mathfrak{n}) - 1 = 2^t u$, $(u, 2) = 1$. Тогда $\mathfrak{n}$ – простой идеал тогда и только тогда, когда для любого $a \in I_{R/\mathfrak{n}}$, $a^u \not\equiv 1 (\mathrm{mod}\ \mathfrak{n})$ существует $k \in \{0, \dots, t-1\}$, такое что $a^{2^k u} \equiv -1 (\mathrm{mod}\ \mathfrak{n})$.

Пусть кольцо R факториальное и удовлетворяет условию А. Тогда $\mathfrak{n}$ – простой идеал тогда и только тогда, когда для любого $a \in I_{R/\mathfrak{n}}$, $Nm(a) \leq f_R\big(Nm(a)\big)$, $(a, \mathfrak{n}) = 1$, $a^u \not\equiv 1 (\mathrm{mod}\ \mathfrak{n})$ существует $k \in \{0, \dots, t-1\}$, такое что $a^{2^k u} \equiv -1 (\mathrm{mod}\ \mathfrak{n})$.

Доказательство. Докажем первую часть теоремы. Предположим, что $\mathfrak{n}$ – простой идеал. Рассмотрим произвольный $a \in I_{R/\mathfrak{n}}$, $a^u \not\equiv 1 (\mathrm{mod}\ \mathfrak{n})$. Из теоремы Эйлера следует, что $a^{2^t u} = a^{\varphi(\mathfrak{n})} \equiv 1 (\mathrm{mod}\ \mathfrak{n})$. Раскладываем на множители и получаем, что выполнено $(a^u - 1)(a^u + 1)(a^{2u} + 1) \dots \left(a^{2^{t-1} u} + 1\right) \equiv 0 (\mathrm{mod}\ \mathfrak{n})$. Из того, что $a^u \not\equiv 1 (\mathrm{mod}\ \mathfrak{n})$ следует, что $a^{2^k u} + 1 \equiv 0 (\mathrm{mod}\ \mathfrak{n})$ для некоторого $k \in \{0, \dots, t-1\}$. Это завершает доказательство необходимости.

Предположим, что $\mathfrak{n}$ – не простой идеал. Пусть $\mathfrak{n}$ раскладывается в произведение простых идеалов следующим образом $\mathfrak{n} = \prod_{i=1}^{r} \mathfrak{p}_i^{\alpha_i}$. Так как норма простого идеала примарная, то обозначим $Nm(\mathfrak{p}_i) = q_i^{f_i}$, где q_i – простой в $\mathbb{Z}$.

Пусть существует такой $j \in \{1, \dots, r\}$, что $\alpha_j > 1$ в разложении $\mathfrak{n}$ на множители. Из теоремы Коши для групп и свойств функции Эйлера следует, что существует $a \in I_{R/\mathfrak{n}}$ порядка q_j. Так как $u \not\equiv 0 \big(mod\ q_j\big)$, то $a^u \not\equiv 1 (mod\ \mathfrak{n})$. Следовательно, существует число $k \in \{0, \dots, t-1\}$, такое что выполнено сравнение $a^{2^k u} \equiv -1 (\mathrm{mod}\ \mathfrak{n})$. Тогда $a^{2^{k+1} u} \equiv 1 (\mathrm{mod}\ \mathfrak{n})$. Значит выполнено $2^{k+1} u \equiv 0 \big(mod\ q_j\big)$. Из последнего сравнения следует, что $Nm(\mathfrak{n}) - 1 \equiv 0 \big(mod\ q_j\big)$, что невозможно.

Следовательно, $\alpha_j = 1$ для любого $j \in \{1, \dots, r\}$. Так как $\mathfrak{n}$ – составное, то $r \geq 2$. Из аналога Китайской теоремы об остатках и того, что элемент -1 имеет порядок 2 в каждой группе $I_{R/\mathfrak{p}_j}$ следует, что существует по крайней мере $2^r - 1 \geq 3$ элемента $I_{R/\mathfrak{n}}$ порядка 2. Пусть $a \not\equiv \pm 1 (mod\ \mathfrak{n})$ является произвольным элементом порядка 2 в группе $I_{R/\mathfrak{n}}$. Из того, что $(u, 2) = 1$ следует, что $a^u \equiv a \not\equiv \pm 1 (mod\ \mathfrak{n})$. Таким образом, существует

$k \in \{0, \dots, t-1\}$, такое что верно $a^{2^k u} \equiv -1 (\mathrm{mod}\, \mathfrak{n})$. Это противоречит тому, что порядок a равен 2. Это завершает доказательство достаточности.

Теперь докажем вторую часть теоремы. Необходимость следует из доказанного ранее. Предположим, что $\mathfrak{n} \in R^* \setminus I_R$ составной идеал нечетной нормы и для любого $a \in I_{R/\mathfrak{n}}$, $Nm(a) \leq f_R\big(Nm(a)\big)$, $(a, \mathfrak{n}) = 1$, $a^u \not\equiv 1 (\mathrm{mod}\, \mathfrak{n})$ существует $k \in \{0, \dots, t-1\}$, такое что $a^{2^k u} \equiv -1 (\mathrm{mod}\, \mathfrak{n})$.

Предположим, что существует такой простой идеал $\mathfrak{p}$, что $\mathfrak{p}^2 | \mathfrak{n}$. Рассмотрим такое отображение $\chi: I_{R/\mathfrak{p}^2} \to I_{R/\mathfrak{p}^2}$, что для всех $a \in I_{R/\mathfrak{p}^2}$ выполнено $\chi(a) = a^{Nm(\mathfrak{p})-1}$. Из предложения 2 следует, что это нетривиальный гомоморфизм. Тогда, из предложения 1, получаем, что существует такой элемент $a \in I_{R/\mathfrak{p}^2}$, что $a^{Nm(\mathfrak{p})-1} \not\equiv 1 (mod\ \mathfrak{p}^2)$.

Предположим, что $a^{Nm(\mathfrak{n})-1} \equiv 1 (mod\ \mathfrak{n})$. Тогда $a^{Nm(\mathfrak{n})-1} \equiv 1 (mod\ \mathfrak{p}^2)$. Тогда $ord_{R/\mathfrak{p}^2}(a) | Nm(\mathfrak{n}) - 1$ и $ord_{R/\mathfrak{p}^2}(a) | \varphi(\mathfrak{p}^2)$. Из этого следует, что $ord_{R/\mathfrak{p}^2}(a) | Nm(\mathfrak{p}) - 1$. Это противоречит выражению $a^{Nm(\mathfrak{p})-1} \not\equiv 1 (mod\ \mathfrak{p}^2)$.

Следовательно, не существует такой простой идеал $\mathfrak{p}$, что $\mathfrak{p}^2 | \mathfrak{n}$. Пусть $\mathfrak{p}$ и $\mathfrak{q}$ различные простые делители $\mathfrak{n}$. Обозначим $v_2(n)$ максимальную степень двойки, делящую $n \in \mathbb{Z}$. Не нарушая общности, предположим, что $v_2(Nm(\mathfrak{p}) - 1) \geq v_2(Nm(\mathfrak{q}) - 1)$. Пусть

$$\mathfrak{d} = \begin{cases} \mathfrak{pq}, \text{если } v_2(Nm(\mathfrak{p}) - 1) = v_2(Nm(\mathfrak{q}) - 1), \\ \mathfrak{p}, \text{если } v_2(Nm(\mathfrak{p}) - 1) > v_2(Nm(\mathfrak{q}) - 1). \end{cases}$$

Рассмотрим такое отображение $\xi: I_{R/\mathfrak{n}} \to I_{R/\mathfrak{n}}$, что для всех $a \in I_{R/\mathfrak{n}}$ выполнено $\xi(a) = \left(\frac{a}{\mathfrak{d}}\right)$. Это отображение является нетривиальным гомоморфизмом. Тогда, из предложения 1, получаем, что существует такой элемент $a \in I_{R/\mathfrak{n}}$, что $\left(\frac{a}{\mathfrak{d}}\right) \not\equiv 1 (mod\ \mathfrak{n})$.

Положим $b = a^u$. Тогда $\left(\frac{b}{\mathfrak{d}}\right) = -1$. Следовательно, $b \not\equiv 1 (mod\ \mathfrak{d})$. Пусть $j \in \mathbb{Z}$ минимальное число, для которого $a^{2^j u} \equiv -1 (mod\ \mathfrak{n})$. Тогда $ord_{R/\mathfrak{p}}(b) = ord_{R/\mathfrak{q}}(b) = 2^{j+1}$.

Рассмотрим два случая. Пусть $v_2(Nm(\mathfrak{p}) - 1) > v_2(Nm(\mathfrak{q}) - 1)$. Тогда $ord_{R/\mathfrak{q}}(b) = 2^{j+1} | \varphi(\mathfrak{q}) = Nm(\mathfrak{q}) - 1$. Следовательно, $ord_{R/\mathfrak{p}}(b) = 2^{j+1} | (Nm(\mathfrak{p}) - 1)/2$. Получаем, что $\left(\frac{b}{\mathfrak{d}}\right) = \left(\frac{a}{\mathfrak{p}}\right) = -1$ и $b^{(Nm(\mathfrak{p})-1)/2} \equiv 1 (mod\ \mathfrak{p})$. Это противоречит критерию Эйлера.

Пусть $v_2(Nm(\mathfrak{p}) - 1) = v_2(Nm(\mathfrak{q}) - 1)$. Тогда $\left(\frac{b}{\mathfrak{d}}\right) = \left(\frac{b}{\mathfrak{p}}\right)\left(\frac{b}{\mathfrak{q}}\right) = -1$. Следовательно, один из множителей равен -1. Пусть $\left(\frac{b}{\mathfrak{p}}\right) = -1$ и $\left(\frac{b}{\mathfrak{q}}\right) = 1$.

Из критерия Эйлера следует, что $b^{(Nm(\mathfrak{q})-1)/2} \equiv 1(mod\ \mathfrak{q})$ и $ord_{R/\mathfrak{p}}(b) = ord_{R/\mathfrak{q}}(b) | (Nm(\mathfrak{q})-1)/2$. Тогда $ord_{R/\mathfrak{p}}(b) | (Nm(\mathfrak{p})-1)/2$. Следовательно, $b^{(Nm(\mathfrak{p})-1)/2} \equiv 1(mod\ \mathfrak{p})$, что противоречит предположению $\left(\frac{b}{\mathfrak{p}}\right) = -1$.

Аналог алгоритма Миллера-Рабина. Пусть дан идеал $\mathfrak{n} \subset R$. Необходимо определить является ли он простым.

1. Найти $\mathfrak{u}, t \in \mathbb{N}$, что $Nm(\mathfrak{n}) - 1 = 2^t u$ и $(2, u) = 1$;
2. Выбрать случайный $a \in I_{R/\mathfrak{n}}$ и вычислить $r_0 \equiv a^u (mod\ \mathfrak{n})$;
3. Если $r_0 = 1$, то вернуть "неизвестно" и завершить алгоритм;
4. Для k от 0 до t выполнить:
   a. Если $r_k = -1$, то вернуть "неизвестно" и завершить алгоритм;
   b. Вычислить $r_{k+1} \equiv r_k^2 (mod\ \mathfrak{n})$;
5. Вернуть "$\mathfrak{n}$ не простой" и завершить алгоритм.

Замечание. Аналог алгоритма Миллера-Рабина является вероятностным. Если был получен ответ "неизвестно", то можно выполнить алгоритм еще раз.

Библиографические ссылки

1. Miller G. Riemann’s Hypothesis and Tests for Primality // Journal of Computer and System Sciences. 1976. № 13(3). P. 300–317.
2. Rabin M.O. Probabilistic Algorithm for Testing Primality // Journal of number theory. 1980. № 12. P. 128–138.
3. Ankeny N.C. The least quadratic non-residue // Ann. of Math. 1952. P. 65–72.
4. Dandan Huang, Yingpu Deng Algorithm for computing the factor ring of an ideal in Dedekind domain with finite rank. Science China Mathematics. 2017. № 61. P. 783-796.
5. Vaskouski M., Kondratyonok N., Prochorov N. Primes in quadratic unique factorization domains // Journal of Number Theory. 2016. № 168. P. 101–116.
6. Прохоров Н.П. Вероятностный и детерминированный аналоги алгоритма Миллера-Рабина для идеалов колец целых алгебраических элементов конечных расширений поля Q // Известия Национальной академии наук Беларуси. Серия физико-математических наук. 2020. № 56. С. 144–156.
7. Petukhova K.A., Tronin S.N. RSA Cryptosystem for Dedekind Rings // Lobachevskii Journal of Mathematics. 2016. № 37. P. 284–287.
8. Bach E. Explicit bounds for primality testing and related problems // Mathematics of Computation. 1990. P. 355–380.

УДК 004.056.5;004.75

КРИПТОГРАФИЧЕСКИЕ МЕТОДЫ ЗАЩИТЫ ИНФОРМАЦИИ, ИСПОЛЬЗУЕМЫЕ В ТЕХНОЛОГИИ БЛОКЧЕЙН

А.Н. Гайдук

Белорусский государственный университет,
пр. Независимости, 4, 220030, г. Минск, Беларусь, gaidukan@bsu.by

Представлены криптографические методы защиты информации, используемые в технологии блокчейн, являющиеся фундаментальной основой его безопасности. Указаны перспективные направления криптографических исследований в области технологии блокчейн.

***Ключевые слова*:** Блокчейн; криптография; криптографические методы защиты информации.

CRYPTOGRAPHIC METHODS OF INFORMATION PROTECTION USED IN BLOCKCHAIN TECHNOLOGY

A.N. Gaiduk

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus,
gaidukan@bsu.by

Cryptographic methods of information protection used in blockchain technology, which are the fundamental basis of its security, are presented. Promising areas of cryptographic research in the field of blockchain technology are indicated.

***Keywords*:** Blockchain; cryptography; cryptographic methods of information protection.

Введение

С момента своего появления технология блокчейн привлекла большое внимание как со стороны научных кругов, правительства, так и бизнеса [1]. Распределенный и децентрализованный характер технологии блокчейна обеспечивает устойчивый к несанкционированному доступу контроль над общим регистром данных. Сегодня технология блокчейн находит широкий спектр применения в различных областях, помимо криптовалют (см., например, [2]). При этом возникает много исследовательских проблем касающихся как безопасности технологии блокчейн, так и ее масштабируемости, и эффективности. Эти проблемы возникают из-за структуры сети и лежащих в ее основе механизмов консенсуса, а также используемых криптографических методов защиты информации.

Поскольку криптография – это обширная область исследований, всегда есть возможность найти новые криптографические методы, чтобы улучшить существующие решения в технологии блокчейн. В данной работе выделены основные методы криптографической защиты информации, используемые в технологии блокчейн, а также указаны перспективные направления криптографических исследований в области технологии блокчейн.

1. Методология исследования / теоретические основы

В последние годы количество публикаций, связанных с технологией блокчейна стремительно растет. Для того чтобы определить множество допустимых статей для анализа была использована методология исследования, которая определяет стратегию поиска соответствующих публикаций, процедуру первичного отбора, критерии включения и исключения, и метод сбора данных для накопления соответствующих публикаций. Чтобы найти соответствующие публикации для нашего исследования, были использованы следующие строки поиска (cryptography) AND (blockchain), (криптография) AND (блокчейн). Поиск осуществлялся через поисковые системы Google и Яндекс. Также проводился поиск в базах данных: 1) Архив eprint IACR, 2) IEEE Xplore, 3) ACM Digital Library 4) ScienceDirect 5) Springer Link.

После процедуры первичного отбора к публикациям применялись критерии включения и исключения, которые определяли соответствие публикации данному исследованию.

2. Результаты и их обсуждение

Структура данных блокчейна представляет собой связанный список, в котором в качестве указателя используется хэш-указатель. Структурной единицей такого списка является блок, который содержит «хэш-указатель» на предыдущий блок и некоторые «данные». Последовательность таких блоков образует цепочку (список), где каждый блок содержит хэш-указатель на предыдущий блок. Хэш-значение предыдущего заголовка блока включается в следующий блок в качестве ссылки, и поэтому изменение даже одного символа в одной из транзакций сделает ссылку недействительной. Данная цепочка (список) и называется блокчейном. Основное различие между блокчейном и связанным списком заключается в том, что ссылки в блокчейне криптографически защищены. Напротив, указатели в связанном списке могут быть изменены в любое время без нарушения целостности данных и, следовательно, может быть изменен по-

рядок следования записей в связанном списке. В блокчейне защищенные ссылки устанавливают порядок следования блоков друг за другом и фактически делают блокчейн структурой данных только для добавления в конец цепочки (списка), т.е. новые данные могут быть добавлены только с новыми блоками. Первый или начальный блок называется блоком генезиса.

Основными криптографическими методами защиты информации, используемыми в технологии блокчейн, являются: хэш-функция, электронная цифровая подпись (ЭЦП), криптографические протоколы доказательства с нулевым разглашением, протоколы конфиденциального вычисления, проверяемая случайная функция. Далее приведем краткое описание применения данных методов криптографической защиты информации в технологии блокчейн.

Определение 1. Хэш-функцией или функцией хэширование называется отображение множества слов произвольной длины в множество слов фиксированной длины:

$$hash: \ \{0,1\}^* \to \{0,1\}^n, \ n \in \mathrm{N}.$$

Используемые в криптографии хэш-функции должны иметь полиномиальную от длины входного слова сложность. В технологии блокчейн хэш-функции используются

- для контроля целостности блоков и транзакций;
- при вычислении адресов участников сети;
- в модели консенсуса;
- при генерации псевдослучайных чисел;
- в алгоритмах ЭЦП.

При конструировании хэш-функций важно учитывать критерии безопасности, которым она должна удовлетворять. Классическими для хэш-функций требованиями являются: стойкость к коллизиям, стойкость к нахождению прообраза, стойкость к нахождению второго прообраза [3].

В технологии блокчейн ЭЦП используются для подписи транзакций, аутентифицируя отправителя и обеспечения контроля целостности транзакции. ЭЦП является одним из наиболее важных криптографических примитивов, благодаря которому блокчейн может быть верифицирован. Наиболее широко используются схемы подписи на основе эллиптических кривых, в частности широко используется алгоритм ЭЦП на эллиптических кривых ECDSA [4]. ECDSA, состоит из трех различных алгоритмов:

- алгоритм генерации ключей;
- алгоритм выработки подписи;
- алгоритма проверка подписи

В блокчейне Биткоина и Ethereum используется эллиптическая кривая известная как кривая Коблица SECP-256k1, которая определена в [5]. В технологии блокчейн также применяются следующие виды подписи:

- мультиподпись;
- слепая подпись;
- подпись на кольце;
- пороговая подпись.

Доказа́тельство с нулевы́м разглаше́нием (информа́ции) в криптографии (англ. Zero-knowledge proof) — интерактивный криптографический протокол, позволяющий одной из взаимодействующих сторон («The verifier» — проверяющей) убедиться в достоверности какого-либо утверждения (обычно математического), не имея при этом никакой другой информации от второй стороны («The prover» — доказывающей). Причём последнее условие является необходимым, так как обычно доказать, что сторона обладает определёнными сведениями в большинстве случаев тривиально, если она имеет право просто раскрыть информацию. Вся сложность состоит в том, чтобы доказать, что у одной из сторон есть информация, не раскрывая её содержание. Доказательства с нулевым разглашением могут быть использованы для обеспечения конфиденциальности данных транзакций в блокчейне. Некоторые блокчейны сети, такие как Zerocoin [6] или Zerocash [7] используют доказательства с нулевым разглашением для того, чтобы транзакции были не отслеживаемыми и их нельзя было соотнести с определенным адресом.

В криптографии протокол конфиденциального вычисления (также безопасное, защищенное или тайное многостороннее вычисление, англ. secure multi-party computation) — криптографический протокол, позволяющий нескольким участникам произвести вычисление, зависящее от тайных входных данных каждого из них, таким образом, чтобы ни один участник не смог получить никакой информации о чужих тайных входных данных. Данный протокол использует блокчейн платформе Enigma [8], в блокчейн Hawki [9] для достижения высокого уровня безопасности, а также позволяет выполнять кроссчейн переводы.

В криптографии проверяемая случайная функция (VRF) — это псевдослучайная функция с открытым ключом, которая обеспечивает доказательства того, что ее выходные данные были вычислены корректно. Владелец секретного ключа может вычислить значение функции, а также соответствующее доказательство для любого входного значения. Используя доказательство и связанный с ним открытый ключ (или ключ проверки), можно проверить, что это значение действительно было рассчитано кор-

ректно, однако эта информация не может быть использована для поиска секретного ключа [10].

Проверяемая случайная функция используется в блокчейнах, основанных на модели консенсуса Proof of Stake [11] для выбора узлов, которые будут публиковать следующий блок и членов комитета по голосованию.

В работе [12] сформулированы следующие исследовательские задачи.

Задача 1. [12] Разработать криптографический протокол, в котором не анонимные пользователи могут публиковать транзакции, которые не могут быть связаны с их сетевыми адресами или другими транзакциями.

Задача 2. [12] Разработать криптографический протокол, в котором не анонимные пользователи могут получать подробную информацию о конкретных транзакциях, не раскрывая, для других участников информацию о том, какие транзакции они ищут.

Задача 3. [12] Разработать эффективные и масштабируемые криптографические протоколы для анонимной публикации в блокчейнах сетях с доступом, требующим разрешения, на основе асинхронных византийскоустойчивых алгоритмов консенсуса для согласования транзакций и процесса смешивания пользовательских данных.

В работе [13] сформулированы следующие исследовательские задачи.

Задача 4. [13] Разработать децентрализованный протокол авторизации для блокчейн сетей с доступом, требующим разрешения, который обеспечит контроль доступа для пользователей.

Задача 5. [13] Разработать модель консенсуса, устойчивой к появлению ветвления в блокчейне.

Задача 6. [13] Разработать устойчивую к кражам блокчейн систему, позволяющую возвращать украденные активы.

Задача 7. [13] Разработать блокчейн систему на основе постквантовой криптографии.

Заключение

Всестороннее исследование базовых криптографических методов защиты информации в технологии блокчейна необходимо для глубокого понимания безопасности и конфиденциальности систем, основанных на технологии блокчейн. В данной работе представлен обзор криптографических методов защиты информации, используемых в технологии блок-

чейн, а также указаны перспективные направления криптографических исследований в области технологии блокчейн.

Библиографические ссылки

1. Paulavi£ius R., Grigaitis S., Igumenov A., and Filatovas E. A decade of blockchain: Review of the current status, challenges, and future directions // Informatica, vol. 30, no. 4, p. 729–748, Jan. 2019.
2. Bodkhe U., Tanwar S., Parekh K., Khanpara P., Tyagi S., Kumar N., and Alazab M. Blockchain for industry 4.0: A comprehensive review // IEEE Access, vol. 8, p. 79764–79800, 2020.
3. Menezes A.J., van Oorschot P.C., Vanstone S.A. Handbook of applied cryptography // CRC Press, 1997.
4. ANSI. Public Key Cryptography for the Financial Services Industry: The Elliptic Curve Digital Signature Algorithm (ECDSA), National Bureau of Standards. ANSI X9.62. 2005. URL: https://standards. globalspec.com/std/1955141/ANSIX9.62.
5. SECG. Recommended Elliptic Curve Domain Parameters. SEC 2 Version 2.0. 2010. URL: https://www.secg.org/sec2-v2.pdf.
6. Miers I., Garman C., Green M., and Rubin A.D. ‘‘Zerocoin: Anonymous distributed e-cash from bitcoin,’’ in Proc. IEEE Symp. Secur. Privacy, May 2013, p. 397–411.
7. Sasson E.B., Chiesa A., Garman C., Green M., Miers I., Tromer E., and Virza M. Zerocash: Decentralized anonymous payments from bitcoin, in Proc. IEEE Symp. Secur. Privacy, May 2014, pp. 459–474.
8. Zyskind G., Nathan O., and Pentland A. Enigma: Decentralized computation platform with guaranteed privacy, 2015, arXiv:1506.03471. [Online]. URL: https://arxiv.org/abs/1506.03471
9. Kosba A., Miller A., Shi E., Wen Z., and Papamanthou C. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts, in Proc. IEEE Symp. Secur. Privacy (SP), May 2016, p. 839–858.
10. Goldberg, Sharon; Vcelak, Jan; Papadopoulos, Dimitrios; Reyzin, Leonid Verifiable Random Functions (VRFs).
11. Blockchain Technology Overview, NIST Technical Series, NISTIR 8202. URL: https://nvlpubs.nist.gov/nistpubs/ir/2018/nist.ir.8202.pdf .
12. Henry R., Herzberg A., and Kate A. Blockchain access privacy: Challenges and directions // IEEE Security Privacy, vol. 16, no. 4, p. 38–45, Jul./Aug. 2018.
13. Raikwar M., Gligoroski D., and Kralevska K. SoK of Used Cryptography in Blockchain // IEEE Access, vol. 7, p. 148550-148575, 2019, doi: 10.1109/ACCESS.2019.2946983.

УДК 517.958:537.8

ЭКРАНИРОВАНИЕ ШИРОКОПОЛОСНЫХ ЭЛЕКТРОМАГНИТНЫХ СИГНАЛОВ МАГНИТОДИЭЛЕКТРИЧЕСКИМ ЭКРАНОМ

В.Т. Ерофеенко[1], О.В. Кравченко [2]

[1]*Учреждение Белорусского государственного университета «НИИ прикладных проблем математики и информатики», пр. Независимости, 4, 220030, г. Минск, Беларусь,* bsu_erofeenko@tut.by

[2]*Научно-технологический центр уникального приборостроения Российской академии наук, Бутлерова д.15, г. Москва, Россия,* ok@ntcup.ru

Рассмотрена краевая задача о проникновении широкополосных импульсных электромагнитных сигналов через плоский однослойный магнито-диэлектрический экран. Представлено аналитическое решение краевой задачи в интегральной форме с помощью монохроматических спектральных электромагнитных полей и финитных атомарных функций, моделирующих сигналы конечной длительности. Вычисляется поле (6), проникающее через экран. Численно проанализирован коэффициент эффективности экранирования временных сигналов, проходящих через экран. Построены графики коэффициента эффективности экранирования.

Ключевые слова: Задача экранирования; плоский экран; широкополосные сигналы; спектральные функции сигналов; атомарные функции; эффективность экранирования; численное моделирование.

SCREENING OF BROAD-BAND ELECTROMAGNETIC SIGNAL BY MAGNETODIELEKTRIC SCREEN

V. T. Erofeenko[a], O. V. Kravchenko[b]

[a]*Belarusian State University «Research Institute for Applied Problems of Mathematics and Informatics», 4 Niezalieznasti Avenue, Minsk 220030, Belarus,* E-mail: bsu_erofeenko@tut.by

[b]*Scientific and Technological Centre of Unique Instrumentation» of the Russian Academy of Sciences, Butlerova 15, Moscow, Russian Federation,* ok@ntcup.ru

Corresponding author: bsu_erofeenko@tut.by

The boundary value problem of the penetration of broad-band pulsed electromagnetic signals through a flat single-layer magnetodielectric screen is considered. An analytical solution of the boundary value problem in integral form is presented utilizing a monochromatic spectral electromagnetic fields and finite atomic functions simulating signals of finite duration. The field (6) penetrating through the screen is calculated. The efficiency coefficient of screening of time signals passing through the screen is numerically analyzed. A graphs of the screening efficiency coefficient is plotted.

Keywords: Screening problem; plane screen; broad-band signals; spectral functions signals; atomic functions; shielding effectiveness; numerical simulation.

Введение

Актуальными для микроэлектроники являются проблемы электромагнитной совместимости технических средств, проблемы защиты научной аппаратуры широкого спектра назначения и информации от воздействий внешних электромагнитных излучений. Для решения этих проблем наиболее эффективным способом является использование электромагнитных защитных экранов и оболочек [1]. Актуальным является разработка методов решения краевых задач экранирования импульсных электромагнитных полей экранами с различными материальными структурами: биизотропными [2], многослойными [2, 3], экранами из пермаллоя [4] и другими.

В предлагаемой работе численно исследовано решение краевой задачи экранирования широкополосных импульсных электромагнитных сигналов [5], представленных через атомарную функцию $\mathrm{up}(x)$ [6], магнитодиэлектрическими экранами. Аналитическое решение задачи получено в интегральном виде и представлено через спектральные базисные электромагнитные поля [7, с.96], сосредоточенные на частотном интервале $-\Omega_{\text{сиг}} < \omega < \Omega_{\text{сиг}}$.

1. Теоретические основы

Сформулируем краевую задачу экранирования импульсных электромагнитных полей магнитодиэлектрическим экраном D (рисунок 1).

Краевая задача. Для заданного первичного поля $\vec{\mathbf{E}}_0, \vec{\mathbf{H}}_0$ требуется определить поля $\vec{\mathbf{E}}_1', \vec{\mathbf{H}}_1'$; $\vec{\mathbf{E}}_2, \vec{\mathbf{H}}_2$; $\vec{\mathbf{E}}, \vec{\mathbf{H}}$, которые удовлетворяют уравнениям

$$\operatorname{rot}\vec{\mathbf{E}}_1' = -\mu_0 \frac{\partial \vec{\mathbf{H}}_1'}{\partial t},\ \operatorname{rot}\vec{\mathbf{H}}_1' = \varepsilon_0 \frac{\partial \vec{\mathbf{E}}_1'}{\partial t} \quad \text{в} \quad D_1, \tag{1}$$

$$\operatorname{rot}\vec{\mathbf{E}}_2 = -\mu_0 \frac{\partial \vec{\mathbf{H}}_2}{\partial t},\ \operatorname{rot}\vec{\mathbf{H}}_2 = \varepsilon_0 \frac{\partial \vec{\mathbf{E}}_2}{\partial t} \quad \text{в} \quad D_2, \tag{2}$$

$$\operatorname{rot}\vec{\mathbf{E}} = -\mu \frac{\partial \vec{\mathbf{H}}}{\partial t},\ \operatorname{rot}\vec{\mathbf{H}} = \varepsilon \frac{\partial \vec{\mathbf{E}}}{\partial t} \quad \text{в} \quad D, \tag{3}$$

граничным условиям непрерывности тангенциальных составляющих полей на плоскостях $z = 0, z = \Delta$:

$$\left(\vec{\mathbf{E}}_{1\tau}-\vec{\mathbf{E}}_{\tau}\right)\Big|_{z=0}=0,\ \left(\vec{\mathbf{H}}_{1\tau}-\vec{\mathbf{H}}_{\tau}\right)\Big|_{z=0}=0,\ \left(\vec{\mathbf{E}}_{2\tau}-\vec{\mathbf{E}}_{\tau}\right)\Big|_{z=\Delta}=0,\left(\vec{\mathbf{H}}_{2\tau}-\vec{\mathbf{H}}_{\tau}\right)\Big|_{z=\Delta}=0;\quad (4)$$

и условиям излучения на бесконечности. ■

В качестве первичного поля $\vec{\mathbf{E}}_0,\vec{\mathbf{H}}_0$, воздействующего на экран, выберем широкополосный электромагнитный сигнал [5], распространяющийся под углом θ_0 к экрану D.

$$\vec{\mathbf{E}}_0(\vec{r},t)=\frac{E_0}{2\pi}\sin(\alpha_+(t))\mathrm{Up}(\alpha_+(t))\vec{e}_y,\ \vec{\mathbf{H}}_0(\vec{r},t)=\frac{E_0}{2\pi Z_0}\sin(\alpha_+(t))\mathrm{Up}(\alpha_+(t))\vec{v}_2^{(-)},\ (5)$$

Решение задачи (1) - (4) определяется формулами

$$\vec{\mathbf{E}}_2(\vec{r},t)=\frac{E_0}{2\pi}U_2(2\alpha_+(t))\vec{e}_y,\ \vec{\mathbf{H}}_2(\vec{r},t)=\frac{E_0}{2\pi Z_0}U_2(2\alpha_+(t))\vec{v}_2^{(-)};$$

$$\vec{\mathbf{E}}_1'(\vec{r},t)=\frac{E_0}{2\pi}U_1(2\alpha_-(t))\vec{e}_y,\ \vec{\mathbf{H}}_1'(\vec{r},t)=\frac{E_0}{2\pi Z_0}U_2(2\alpha_-(t))\vec{v}_2^{(+)}. \quad (6)$$

где

$$U_2(x)=\int_{-1}^{1}\frac{\mathrm{up}_{\text{ш}}^{(-)}(\overline{\omega})}{|D|^2}(\cos(\overline{\omega}\alpha_v)\sin(\overline{\omega}(x-\alpha_0))+B\sin(\overline{\omega}\alpha_v)\cos(\overline{\omega}(x-\alpha_0)))d\,\overline{\omega},$$

$$U_1(x)=\frac{1}{2}\int_{-1}^{1}\frac{\mathrm{up}_{\text{ш}}^{(-)}(\overline{\omega})}{|D|^2}B^{(-)}(\sin(2\,\overline{\omega}\alpha_v)\cos(\overline{\omega}\,x)-B(1-\cos(2\,\overline{\omega}\alpha_v))\sin(\overline{\omega}\,x))d\,\overline{\omega},$$

$$|D|^2=(\cos^2(\alpha_v\overline{\omega})+B^2\sin^2(\alpha_v\overline{\omega}),\ \alpha_v=\Omega_{\text{сиг}}\overline{v}\frac{\Delta}{c}.\ \alpha_0=\Omega_{\text{сиг}}\frac{\Delta}{c}\cos\theta_0,$$

$$B=\frac{1}{2}\left(\frac{\overline{v}}{\mu_r\cos(\theta_0)}+\frac{\mu_r\cos(\theta_0)}{\overline{v}}\right),\ \ B^{(-)}=\frac{1}{2}\left(\frac{\overline{v}}{\mu_r\cos(\theta_0)}-\frac{\mu_r\cos(\theta_0)}{\overline{v}}\right).$$

$$\vec{v}_2^{(\mp)}=\mp\cos\theta_0\,\vec{e}_x+\sin\theta_0\,\vec{e}_z,\quad \overline{v}=\sqrt{\varepsilon_r\mu_r-\sin^2\theta_0},\ \mu=\mu_r\mu_0,\quad \varepsilon=\varepsilon_r\varepsilon_0.$$

Используется атомарная функция $\mathrm{up}(x)$ и спектральная функция $\mathrm{Up}(y)$ [6].

$$\mathrm{up}(x)\neq 0\ \text{при}\ |x|<1,\ \ \mathrm{up}(x)=0\ \text{при}\ |x|\geq 1, \quad (7)$$

$$\mathrm{up}(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{Up}(y)\exp(-ixy)dy,\ \ \mathrm{Up}(y)=\prod_{n=1}^{\infty}\mathrm{sinc}\left(\frac{y}{2^n}\right),\ \mathrm{sinc}(x)=\frac{\sin x}{x},$$

$$\mathrm{up}_{\text{ш}}^{(-)}(\overline{\omega})=\mathrm{up}(2\overline{\omega}-1)-\mathrm{up}(2\overline{\omega}+1).$$

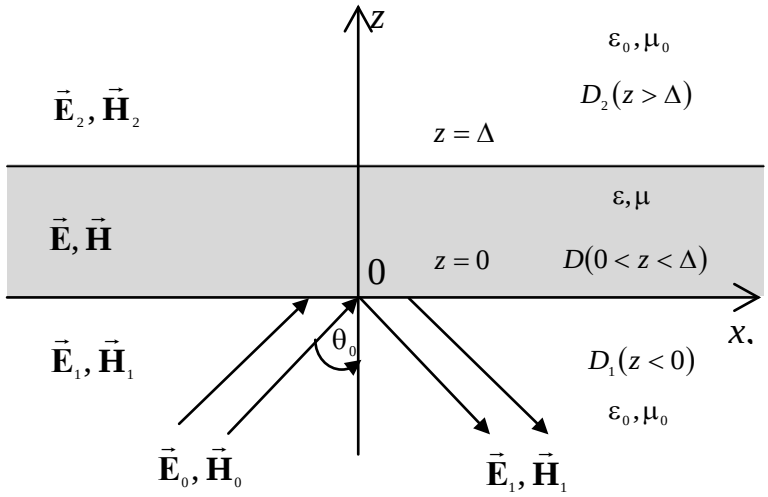


Рисунок 1 – Экранирование широкополосных электромагнитных сигналов магнитодиэлектрическим экраном.

2. Результаты и их обсуждение

Проведено численное моделирование коэффициента эффективности экранирования $Э$ импульсного широкополосного электромагнитного сигнала, показывающего во сколько раз ослабевает сигнал при прохождении через экран.

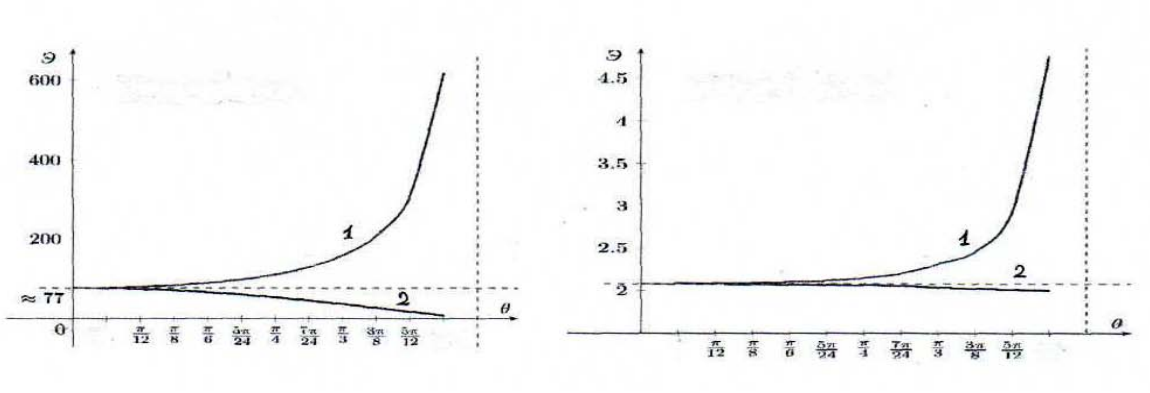
Коэффициент эффективности экранирования определим соотношением

$$Э=\frac{\max\limits_{-\infty<t<\infty}\left|\vec{E}_0(z=0,t)\right|}{\max\limits_{-\infty<t<\infty}\left|\vec{E}_2(z=\Delta,t)\right|},\qquad(8)$$

где поля $\vec{E}_0,\vec{E}_2$ определены в формулах(5), (6).

На рисунке 2 построены графики для коэффициента эффективности экранирования (8) электрического поля широкополосного сигнала для некоторых значений параметров. Коэффициент эффективности (8) зависит от пяти параметров: $\mu_r,\varepsilon_r,\Delta,\Omega_{\text{сиг}},\theta_0$; μ_r – относительная

магнитная проницаемость экрана, ε_r – относительная диэлектрическая проницаемость экрана, Δ – толщина экрана, $-\Omega_{\text{сиг}} < \omega < \Omega_{\text{сиг}}$ частотный интервал широкополосного сигнала, θ_0 – угол падения поля сигнала на экран.



a) b)

Рисунок 2 – Зависимость коэффициента эффективности экранирования от θ_0 при значении параметров: $\Delta = 10^{-3}$ м ; графики:

1) $\Omega_{\text{сиг}} = 10^5$ Гц, **2)** $\Omega_{\text{сиг}} = 10^3$ Гц **;** a) $\mu_r = 10^8, \varepsilon_r = 1$**;** b) $\mu_r = 1, \varepsilon_r = 10^8$.

Разработанная методика может быть использована для решения краевых задач экранирования узкополосных электромагнитных сигналов магнитодиэлектрическим экраном.

Заключение

Показано, что коэффициент эффективности экранирования ТЕ- поляризованного широкополосного электромагнитного сигнала (вектор $\vec{E}_0$ параллелен экрану) принимает большие значения для магнитного экрана $\mu_r = 10^8, \varepsilon_r = 1$(рисунок 2, a)) и малые значения для диэлектрического экрана $\mu_r = 1, \varepsilon_r = 10^8$ (рисунок 2, b)). Показано, что уменьшение частоты $\Omega_{\text{сиг}}$ сигнала приводит к резкому уменьшению эффективности экранирования.

Библиографические ссылки

1. Аполлонский С. М. Расчёт электромагнитных экранирующих оболочек. Ленинград: Энергоиздат, 1982. 144 с.

2. Ерофеенко В.Т., Малый С.В. Алгоритм численного исследования экранирующих свойств многослойных экранов из композитных материалов // Информатика. 2010. № 37(4). С. 96–104.
3. Ерофеенко В.Т., Бондаренко В.Ф. Экранирование магнитного импульса пленочным многослойным экраном с чередующимися магнитными и немагнитными слоями // Журнал технической физики. 2017. № 87(6). С. 831–836.
4. Ерофеенко В.Т., Громыко Г.Ф., Заяц Г.М. Численное моделирование задач экранирования импульсных электромагнитных полей экранами из пермаллоя // Дифференциальные уравнения. 2021. № 57 (12). С. 1682–1697.
5. Ерофеенко В.Т., Урбанович А.И. Конструирование импульсных широкополосных и узкополосных электромагнитных сигналов, распространяющихся в пространстве, с применением атомарных функций // Физические основы приборостроения. 2021. № 10 (2). С. 34–41.
6. Кравченко В.Ф. Лекции по теории атомарных функций и некоторым их приложениям. Монография. М.: Радиотехника, 2003. 512 с.
7. Ерофеенко В.Т., Козловская И.С. Аналитическое моделирования в электродинамике. Минск: БГУ, 2014. 304 с.

УДК 534; 53.08:681.3

ОПТИМИЗАЦИОННЫЙ МЕТОД ОЦЕНКИ ЗАЩИЩЕННОСТИ РЕЧЕВОГО СИГНАЛА В КАНАЛЕ УТЕЧКИ ИНФОРМАЦИИ

В.К. Железняк, Е.Р. Адамовский

Учреждение образования «Полоцкий государственный университет имени Евфросинии Полоцкой», ул. Блохина, 29, 211440, г. Новополоцк, Беларусь,
v.zheleznjak@psu.by, e.adamovsky@psu.by

Предложен метод оценки защищенности канала утечки информации на основе взаимно-корреляционного анализа огибающей измерительного сигнала в речевом диапазоне частот. Алгоритм включает генерацию измерительного сигнала и выделение его огибающей, излучение и измерение в канале утечки, выделение огибающей результирующего сигнала, вычисление коэффициента корреляции между исходной и полученной огибающими, сравнение с пороговым значением. Выполнено имитационное моделирование метода программной среде MatLab. Полученные результаты подтверждают большую эффективность использования огибающей по сравнению с исходным сигналом, а также демонстрируют преимущество речевых сигналов перед гармоническими сигналами в качестве измерительных для оценки защищенности канала утечки информации.

Ключевые слова: канал утечки информации; огибающая речевого сигнала; взаимная корреляция; техническая защита информации.

OPTIMIZATION METHOD FOR ESTIMATION OF SPEECH SIGNAL SECURITY IN THE INFORMATION LEAKAGE CHANNEL

V.K. Zheleznyak, Y.R. Adamovskiy

Euphrosyne Polotskaya State University of Polotsk, st. Blokhin, 29, 211440, Novopolotsk, Republic of Belarus
Corresponding author: e.adamovsky@psu.by

A method of information leakage channel security estimating based on the test speech signal envelope cross-correlation analysis is proposed and its includes: test signal generating and extracting its envelope, emitting and measuring in a leakage channel, extracting the resulting signal envelope, calculating the correlation coefficient between the original and received envelopes, and comparing with a threshold value. Simulation modeling of the method in the MatLab software environment has been performed. The obtained results confirm the greater efficiency of using the envelope compared to the original signal, and demonstrate the speech signals advantage over harmonic signals as test signals for assessing the information leakage channel security.

Keywords: information leakage channel; speech signal envelope; cross-correlation; technical information security.

Введение

Актуальность разработки методов оценки защищенности речевых сигналов (РС) в каналах утечки информации (КУИ) заключается в отсутствии единой модели восприятия речи [1, 2].

Питание усилителей, в том числе – аудиосистем, осуществляется через сеть переменного тока. Изменение потребления тока нагрузки приводит к нестабильности по току на входе стабилизатора [3]. Таким путем, РС из питаемой микрофонной системы способен проникать в электромагнитный КУИ в составе излучения усилителя. Другим способом образования КУИ РС является цифро-аналоговое преобразование (ЦАП), порождающее побочные излучения, которые содержат информацию об исходном сигнале [4].

Известно, что РС характеризуется спектром сложной формы в широком диапазоне от 90 до 10-13 кГц [5]. Во временной области для РС может быть вычислена огибающая в узкой и заранее известной полосе низких частот, которая отражает скорость смены фонем речи.

Подобная предсказуемость открывает возможности для повышения точности оценки защищенности РС. Поэтому в данной работе предлагается метод оценки защищенности речевой информации КУИ на основе анализа огибающей измерительного речевого сигнала в точках излучения и наблюдения.

1. Методология исследования

Рассмотрим аналитический сигнал $s(t)$, реальная $s_{re}(t)$ и мнимая $s_{im}(t)$ части которого связаны преобразованием Гильберта [6]. Практическая значимость соотношения (1) заключается в возможности выделения из его частей мгновенной амплитуды $u(t)$ (2), что применимо и к реальным сигналам.

$$s_{im}(t) = \int_{-\infty}^{\infty} s_{re}(\tau)/\pi(t-\tau)d\tau \quad (1)$$

$$u(t) = \sqrt{s_{re}^2(t) + s_{im}^2(t)} \quad (2)$$

Значения мгновенной амплитуды соответствуют понятию огибающей сигнала, которой оперируют при обработке амплитудно-модулированных (АМ) сигналов.

При оценке защищенности РС в КУИ требуется установить взаимосвязь между сигналами в точке излучения и точке наблюдения – в

канале утечки. В качестве меры их схожести предлагается значение коэффициента корреляции Пирсона, который вычисляется согласно формуле (3), и обозначаемого как R [7]. Коэффициент отражает, насколько изменение амплитуды одного сигнала влияет на изменение амплитуды другого сигнала.

$$R = M[(s(t) - M(s(t))) \times (u(t) - M(u(t)))]/(\sigma_{s(t)} \times \sigma_{u(t)}) \qquad (3)$$

где M – математическое ожидание;

σ – среднеквадратическое отклонение.

Предлагается метод оценки защищенности РС КУИ на основе взаимно-корреляционного анализа, заключающийся в генерации и излучении измерительного РС, огибающая которого сравнивается с огибающей в точке наблюдения. Оптимизация метода заключается в использовании сигнала, который был бы наиболее устойчив к шумам в КУИ за счет структурных свойств. Алгоритм (рисунок 1) включает шаги:

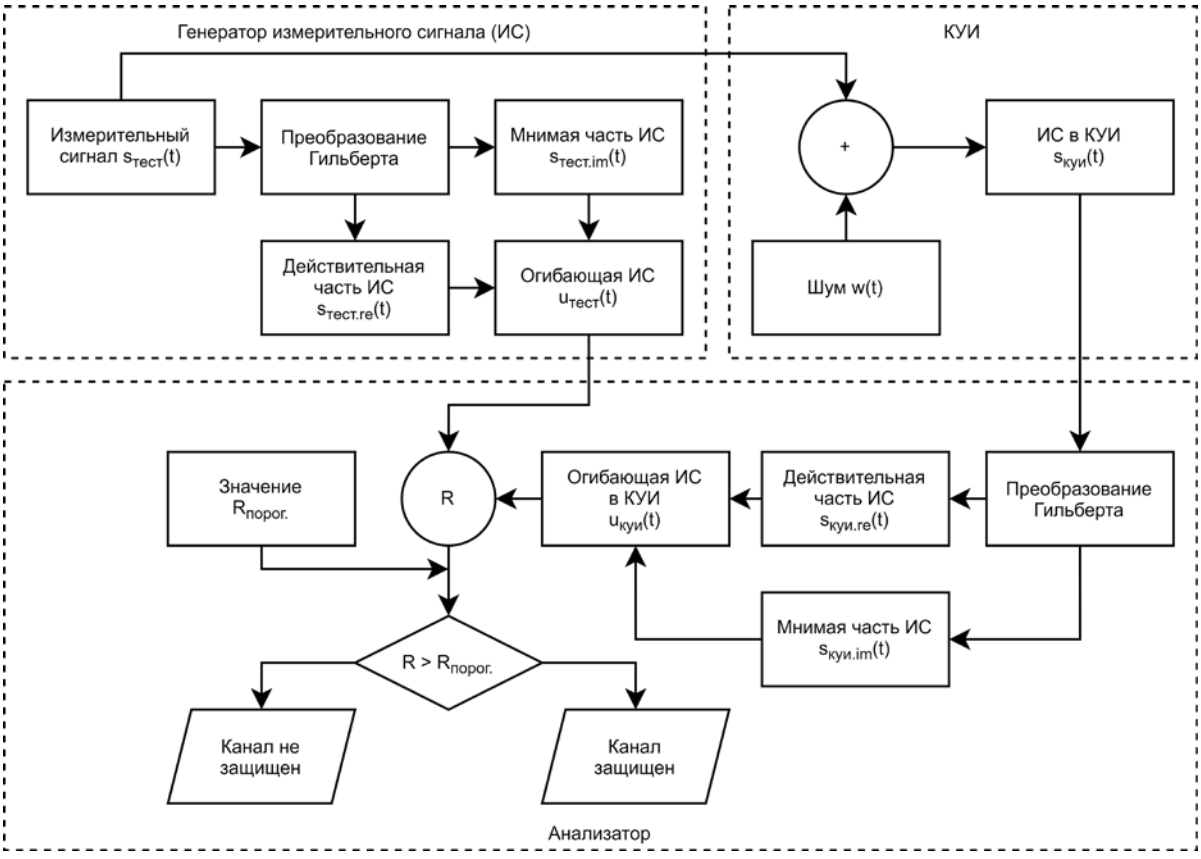


Рисунок 1 – Алгоритм имитационной модели метода оценки

1. Генерация измерительного сигнала $s_{\text{тест}}(t)$ в речевом диапазоне частот в аналитической форме согласно формуле (1).

2. Выделение огибающей $u_{\text{тест}}(t)$ из измерительного сигнала $s_{\text{тест}}(t)$ согласно формуле (2).

3. Излучение измерительного сигнала $s_{\text{тест}}(t)$ в КУИ и его измерение в точке наблюдения как $s_{\text{КУИ}}(t)$, который в простейшей модели КУИ может быть представлен как аддитивная смесь с шумом в КУИ $\omega(t)$.

4. Выделение из $s_{\text{КУИ}}(t)$ огибающей $u_{\text{КУИ}}(t)$ аналогично п. 2.

5. Обработка $u_{\text{тест}}(t)$ и $u_{\text{КУИ}}(t)$ взаимно-корреляционным способом, получение значения коэффициента корреляции R согласно формуле (3).

6. Сравнение полученной величины R с нормативным пороговым значением $R_{\text{порог}}$.

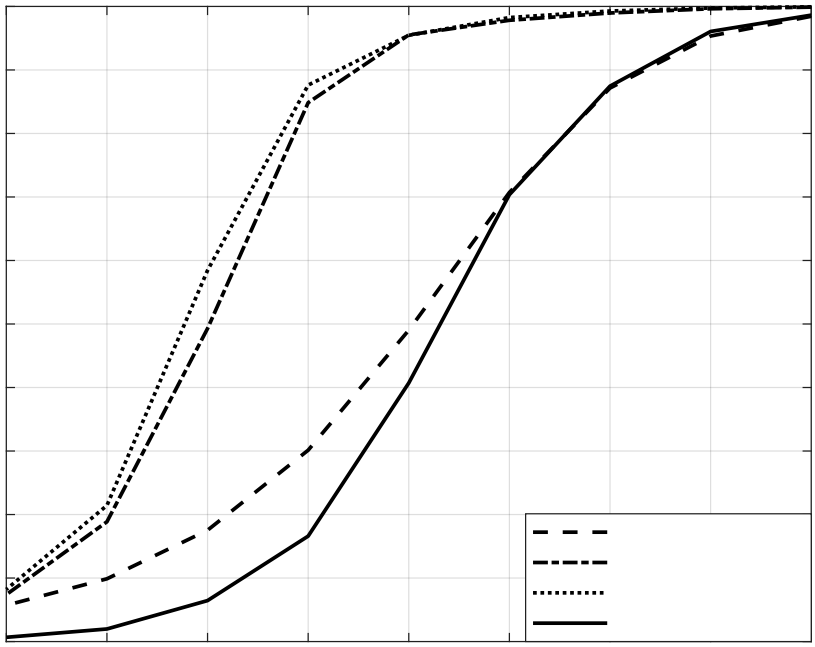
Следует отметить, что результаты взаимно-корреляционного анализа могут быть использованы только при условии синхронизации обоих сигналов относительно друг друга.

2. Результаты и их обсуждение

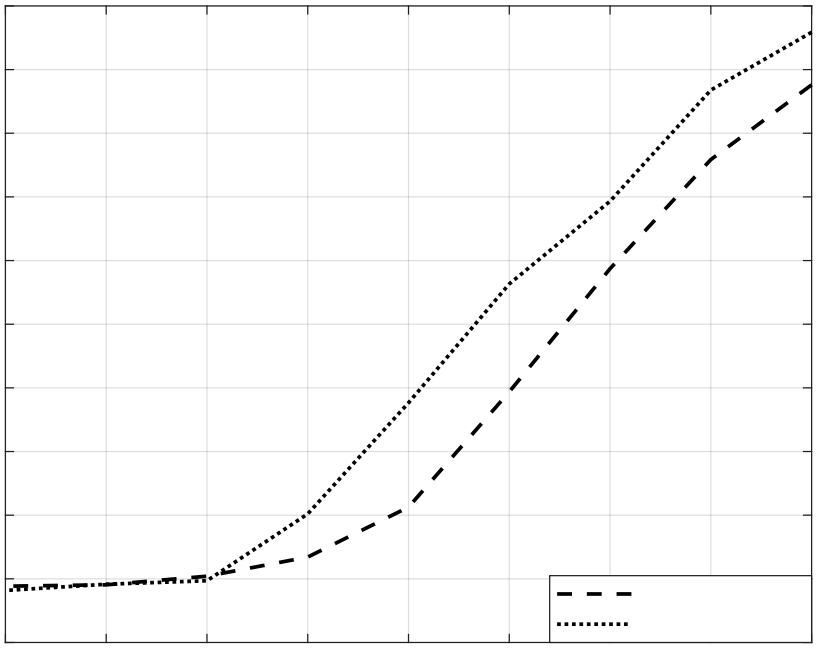
Имитационное моделирование метода реализовано в программной среде MatLab, на рисунке 2 показаны его результаты: значения коэффициентов R и модуляции m измерительных сигналов как среднее арифметическое 10 измерений согласно представленному алгоритму.

В качестве измерительных сигналов был использован РС $s_{\text{речь}}(t)$ – озвученная на русском языке фраза-панграмма; выделенная огибающая $u_{\text{речь}}(t)$; гармонический АМ-сигнал $s_{\text{гарм.АМ}}(t)$ при $m = 1$; выделенная огибающая $u_{\text{гарм.АМ}}(t)$. Исходные РС подвергались зашумлению с различными отношениями сигнал/шум (ОСШ). В результате были получены сигналы $s_{\text{речь.КУИ}}(t)$, $u_{\text{речь.КУИ}}(t)$, $s_{\text{гарм.АМ.КУИ}}(t)$ и $u_{\text{гарм.АМ.КУИ}}(t)$ соответственно. Огибающие ограничивались по частоте до 30 Гц. Измерено соотношение исходного m и полученного $m_{\text{30Гц.КУИ}}$ коэффициента модуляции. Дополнительно для сигналов $u_{\text{речь}}(t)$ и $u_{\text{речь.КУИ}}(t)$ был реализован вариант без ограничения по частоте для исследования влияния высокочастотной (ВЧ) составляющей на результаты моделирования, значение $m_{\text{КУИ}}$.

Из рисунка 2а следует, что оценка корреляционных свойств сигналов во всей доступной частотной полосе дает низкие значения, поскольку влияние широкополосного шума снижает величину схожести. Это подтверждается быстрым спадом не ограниченных по частоте кривых. Рисунок 2б демонстрирует характер падения коэффициента модуляции $m_{\text{30Гц.КУИ}}$, вычисленного по огибающим сигналов.



а)



б)

Рисунок 2 – Результаты имитационного моделирования метода при различных уровнях шума, сравнение полученных значений: а) коэффициентов взаимной корреляции R; б) коэффициентов модуляции m

Показано, что модуляция речевого сигнала более устойчива к шуму, чем гармонического модулированного сигнала.

Заключение

Представлен метод оценки защищенности канала утечки информации на основе взаимно-корреляционного анализа огибающей измерительного сигнала в речевом диапазоне частот и результаты имитационного моделирования метода. Произведен сравнительный анализ результатов для огибающей речевого сигнала, исходного речевого сигнала и гармонического амплитудно-модулированного сигнала Показаны преимущества использования огибающей речевого сигнала для оценки защищенности канала утечки информации, поскольку в этом случае достигаются большие значения коэффициента корреляции в шумах высокого уровня, следовательно использование огибающей РС оптимально.

Библиографические ссылки

1. Анохин В.В., Герасименко Е.А., Кондратьев А.В. Рассмотрение критериев защищённости речи на основе словесной и смысловой разборчивости // Специальная техника. 2016. № 6. С. 22–28.
2. Железняк В. К., Макаров Ю. К., Хорев А. А. Некоторые методические подходы к оценке эффективности защиты речевой информации // Специальная техника. 2000. № 4. С. 39.
3. Костиков В.Г., Парфенов Е.М., Шахнов В.А. Источники электропитания электронных средств. Схемотехника и конструирование // Горячая линия–Телеком. 2001. С. 344.
4. Адамовский Е.Р. Излучение цифро-аналогового преобразователя при обработке тестовых сигналов // Современные средства связи: материалы XXVI Междунар. науч.-техн. конф. Минск: Белорусская государственная академия связи. 2021. С. 124–126.
5. Трушин В.А., Иванов А.В., Рева И.Л. О корректировке методики оценки защищённости речевой информации от утечки по техническим каналам // Специальная техника. 2016. № 6. С. 22–30.
6. Бутырский Е.Ю. Преобразование гильберта и его обобщение // Научное приборостроение. 2014. № 24(4). С. 30–37.
7. Железняк В. К. Защита информации от утечки по техническим каналам: учебное пособие. М.: ГУАП, 2006. 187 с.

УДК 004.056.5

КОНФИДЕНЦИАЛЬНОСТЬ ВЫБОРА В СИСТЕМАХ ЭЛЕКТРОННОГО ГОЛОСОВАНИЯ

М.А. Казловский

Белорусский государственный университет,
пр. Независимости 4, 220030, г. Минск, Беларусь, mkazl@yahoo.com

В работе рассматривается проблема обеспечения конфиденциальности выбора в системах электронного голосования. Выполнен аналитический обзор криптографических механизмов, обеспечивающих выполнение конфиденциальности выбора. Построены ограничения на использование сочетаний данных механизмов, в результате чего получены оптимальные наборы. Приведены примеры использования этих наборов в реальных системах электронного голосования.

***Ключевые слова*:** Электронное голосование; конфиденциальность выбора; доказательство с нулевым разглашением; mix сети; гомоморфное шифрование.

FAIRNESS IN ELECTRONIC VOTING SYSTEMS

M. A. Kazlouski

Belarusian State University, 4 Nezavisimosti Avenue, Minsk, 220030, Belarus,
mkazl@yahoo.com

The paper deals with the problem of ensuring the fairness in electronic voting systems. An analytical review of cryptographic mechanisms that ensure the fairness is carried out. Restrictions on the use of combinations of these mechanisms are built, as a result, optimal sets are obtained. Examples of the use of these sets in real electronic voting systems are given.

***Keywords*:** Electronic voting; fairness; zero-knowledge proof; mix networks; homomorphic encryption.

Введение

В процессе голосования избиратель, как правило, получает бюллетень, заполняет его и опускают в урну для голосования. Не нарушая общности, можно считать, что до завершения процедуры голосования никто не имеет доступа к урне и, соответственно, не может узнать промежуточные итоги голосования. Данное требование позволяет обеспечить объективность голосования: избиратели не смогут учитывать текущие результаты при принятии решений о своем участии в выборах и схеме заполнения бюллетеня, а заинтересованные лица не будут иметь соблазна осуще-

ствить подкуп избирателей, если, согласно текущим результатам, для победы нужного им варианта не хватает сравнительно небольшого числа голосов.

В случае проведения электронного голосования такой подход может не сработать, так как обычно сразу после публикации бюллетеня доступ к его содержимому имеют все участники избирательного процесса (избиратели, избирательная комиссия, наблюдатели). Таким образом, возникает необходимость в защите содержимого бюллетеня, которая обеспечит его конфиденциальность до завершения процедуры голосования. На английском языке такое свойство протокола электронного голосования называется «fairness», в качестве русского аналога мы будем использовать словосочетание «конфиденциальность выбора».

В научных статьях, посвященным децентрализованным системам электронного голосования, вопрос соблюдения конфиденциальности выбора обычно либо не поднимается вовсе, либо свойство выполняется неявно (например, если система голосования обеспечивает защиту от принуждения, то обязательно будет достигаться и конфиденциальность выбора). Данная работа рассматривает системы электронного голосования в контексте соблюдения указанного свойства. Выделены криптографические механизмы, которые могут использоваться для достижения конфиденциальности выбора, изучена их сочетаемость и установлены оптимальные сочетания. Найденные наборы соотнесены с известными системами электронного голосования.

1. Криптографические механизмы при обеспечении конфиденциальности выбора

Единственным возможным вариантом защиты содержимого бюллетеня является его шифрование. Однако организовать шифрование можно используя различные криптографические механизмы. Рассмотрим криптографические механизмы, между вариантами реализации которых необходимо выбирать при проектировании системы обеспечении конфиденциальности выбора.

1) *Выполняющий расшифрование субъект*. Расшифрование бюллетеня может осуществляться:

a. избирателем – используется симметричный или ассиметричный алгоритм шифрования с публикацией после завершения голосования секретного или личного ключа соответственно;

b. избирательной комиссией – используется асимметричный алгоритм шифрования: открытый ключ публикуется до начала голосования

и используется избирателем для зашифрования, а личный ключ используется избирательной комиссией для расшифрования.

2) *Формат публикации бюллетеня.* Избиратель всегда публикует зашифрованный бюллетень. При этом:

a. ничего дополнительно не публикуется;

b. дополнительно публикуются неинтерактивные доказательства с нулевым разглашением, подтверждающие корректность зашифрованного значения и процедуры зашифрования.

3) *Подтверждение корректности расшифрования.* Если расшифрование выполняется избирательной комиссией, то подтверждение корректности расшифрования может осуществляться путем:

a. публикации личного ключа – любой желающий может проверить результаты расшифрования;

b. публикации неинтерактивных доказательств с нулевым разглашением для расшифрованного бюллетеня или суммы бюллетеней – любой желающий может проверить корректность опубликованных доказательств.

4) *Подсчет результатов.* Если расшифрование выполняется избирательной комиссией и алгоритм шифрования обладает гомоморфизмом, то подсчет результатов может осуществляться путем:

a. суммирования результатов в расшифрованных по отдельности бюллетенях;

b. суммирования бюллетеней в зашифрованном виде с последующим расшифрованием итогового результата голосования.

5) *Обработка бюллетеней.* Если расшифрование выполняется избирательной комиссией, то при обработке бюллетеней:

a. ничего дополнительно не выполняется;

b. избирательная комиссия осуществляет перемешивание бюллетеней с помощью mix сетей.

6) *Организация избирательной комиссии.* Если расшифрование выполняется избирательной комиссией, то она может организовываться различными способами:

a. централизованная комиссия, личный ключ которой разделен на несколько частичных секретов;

b. децентрализованная комиссия, в которой у каждого члена комиссии есть свой личный ключ.

2. Оптимальные сочетания криптографических механизмов при обеспечении конфиденциальности выбора

Комбинируя различные варианты описанных выше криптографических механизмов, можно получить $2^6 = 64$ сочетания. Но часть этих сочетаний невозможна, а часть не имеет смысла. Сформулируем правила, позволяющие снизить число рассматриваемых сочетаний:

- Если выбран вариант 1.a, то механизмы 3 – 6 не имеют смысла, т.к. избирательная комиссия не участвует в расшифровании бюллетеней.
- Если выбран вариант 1.a или вариант 3.a, то нет смысла выбирать вариант 2.b, т.к. в данном случае нет необходимости блокировать некорректные бюллетени на этапе голосования.
- Если выбран вариант 3.a, то нет смысла выбирать вариант 4.b, т.к.
  в случае раскрытия личного ключа избирательной комиссии любой желающий сможет расшифровать все бюллетени по отдельности.
- Если выбран вариант 4.b, то должен использоваться вариант 2.b, т.к. иначе избиратель может зашифровать некорректные данные и подсчет голосов будет проведен некорректно.
- Если выбран вариант 5.b, то обязательно должен быть выбран вариант 2.b, т.к. для использования mix сетей необходимо, чтобы бюллетень имел корректный формат.
- Если выбран вариант 5.b, то обязательно должен быть выбран вариант 3.b, т.к. если проводимые избирательной комиссией перемешивания не будут сопровождаться доказательствами, то возможно нарушение принципа состоятельности голосования.
- Если выбран вариант 5.b, то нет смысла выбирать вариант 4.b, т.к.
  получение итогового результата с помощью гомоморфного шифрования после использования mix сетей не дает дополнительных гарантий безопасности.
- Если выбран вариант 5.b, то обязательно должен быть выбран вариант 6.b, т.к. использование mix сетей предполагает перемешивание бюллетеней каждым членом комиссии, что невозможно для централизованной комиссии.
- Если выбран вариант 6.a, то нет смысла выбирать вариант 3.b, т.к. доказательства с нулевым разглашением нужны лишь при

частичном расшифровании бюллетеня, что характерно для децентрализованной комиссии.

После применения сформулированных выше ограничений к списку из 64 возможных сочетаний криптографических механизмов, получим семь возможных наборов сочетаний: усеченный набор [1.a, 2.a], который обозначим как (0) и ряд полных наборов:

- [1.b, 2.a, 3.a, 4.a, 5.a, 6.a] – (1);
- [1.b, 2.a, 3.a, 4.a, 5.a, 6.b] – (2);
- [1.b, 2.a, 3.b, 4.a, 5.a, 6.b] – (3);
- [1.b, 2.b, 3.b, 4.a, 5.a, 6.b] – (4);
- [1.b, 2.b, 3.b, 4.a, 5.b, 6.b] – (5);
- [1.b, 2.b, 3.b, 4.b, 5.a, 6.b] – (6).

Разделим найденные наборы на классы. По типу субъекта, выполняющего расшифрование голосов, можно выделить три класса. Класс A – расшифрование выполняется пользователем, в который входит набор (0), класс B – расшифрование выполняется централизованной избирательной комиссией, в который входит набор (1) и класс C – расшифрование выполняется децентрализованной избирательной комиссией, в который входят наборы (2) – (6). В свою очередь класс C можно разделить на 3 подкласса: подкласс C1 – протокол голосования использует mix сети, в который входит набор (5), подкласс C2 – протокол голосования использует гомоморфное шифрование, в который входит набор (6) и подкласс C3 – протокол голосования не использует дополнительных криптографических механизмов в процессе обработки зашифрованных голосов, в который входят наборы (2) – (4).

Таким образом, все классы и подклассы, кроме подкласса C3 содержат по одному набору. Подкласс C3 состоит из трех наборов, при этом с ростом номера набора наблюдается увеличение вычислительной сложности протокола голосования. Так набор (3) отличается от набора (2) тем, что комиссия не раскрывает ключ расшифрования, при этом предоставляя доказательства корректности расшифрования. А набор (4) отличается от набора (3) тем, что избиратель дополнительно строит доказательства корректности содержимого бюллетеня. Отметим, что в контексте свойства конфиденциальность выбора такие «усиления» видятся избыточными, так как добавляют дополнительную вычислительную нагрузку при формировании и проверке доказательств, не предоставляя при этом дополнительных гарантий безопасности. Однако, возможно, они могут быть полезны для поддержания каких-либо других свойств протоколов электронного голосования.

Среди популярных систем электронного голосования большинство входит или в подкласс C1 (например, JCJ [1], Civitas [2]), или в подкласс C2 (например, Cobra [3], система голосования на выборах в РФ [4]). Это связно с тем, что протокол электронного голосования должен поддерживать не только конфиденциальность выбора, но и ряд других криптографических свойств, таких как анонимность и защита от принуждения, для выполнения которых и используются такие криптографические механизмы как mix сети и гомоморфное шифрование.

Таким образом, получена классификация криптографических механизмов, используемых для соблюдения свойства конфиденциальности выбора, которую можно использовать при проектировании и реализации систем электронного голосования, обладающих данным свойством.

Библиографические ссылки

1. Juels A., Catalano D., Jakobsson M. Coercion-resistant electronic elections [Electronic resource]. URL: https://eprint.iacr.org/2002/165.pdf
2. Clarkson M.R., Chong S., Myers A.C. Civitas: toward a secure voting system [Electronic resource]. URL: https://www.cs.cornell.edu/andru/papers/civitas-tr.pdf
3. Essex A., Clark J., Hengartner U. Cobra: toward concurrent ballot authorization for internet voting [Electronic resource]. URL: https://users.encs.concordia.ca/~clark/papers/2012_evt.pdf
4. Программно-технический комплекс, обеспечивающий дистанционное электронное голосование избирателей (участников референдума) вне зависимости от места их нахождения. Описание протокола ДЭГ к выборам, голосование на которых состоится 17, 18 и 19 сентября 2021 г. [Электронный ресурс]. URL: https://deg.rt.ru/landing/materials/7/deg2021_protocol.pdf

УДК 004.056

ОБОБЩЕННЫЙ ПАРАМЕТРИЗОВАННЫЙ АЛГОРИТМ ВОССТАНОВЛЕНИЯ ПРООБРАЗА ХЕШ-ФУНКЦИИ MD4 МЕТОДОМ ПОЛНОГО ОПРОБОВАНИЯ

Н.А. Коновалов

Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский ядерный университет «МИФИ», 115409, Россия, Москва, Каширское шоссе, 31, nikitakonovalov2013@yandex.ru

Данная работа посвящена исследованию криптографической хеш-функции MD4 и некоторых особенностей её конструкции, позволяющих редуцировать функцию для задачи восстановления прообраза по известному образу при некоторых известных характеристиках прообраза методом полного опробования. На основе данных особенностей предложен редуцированный параметризированный алгоритм MD4, значительно сокращающий количество алгоритмических и логических операций. Частным случаем алгоритма является оптимизированный алгоритм поиска прообраза методом полного опробования, предложенный разработчиками специализированного программного обеспечения Hashcat [1]. Обобщенный алгоритм позволяет сократить количество шагов обновления состояния во внутреннем цикле алгоритма на 52% , а также сократить количество наиболее трудоемкой операции сложения по модулю 2^{32} на 66% в лучшем случае.

***Ключевые слова*:** хеш-функция MD4; восстановление прообраза; обратимые преобразования; метод полного опробования; обобщенный параметризованный алгоритм; оптимизация.

GENERALIZED PARAMETERIZED ALGORITHM FOR RECOVERING THE PREIMAGE OF THE MD4 HASH FUNCTION BY THE BRUTE FORCE METHOD

N.A. Konovalov

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), 31 Kashirskoe Shosse, Moscow, 115409, Russian Federation, nikitakonovalov2013@yandex.ru

This work is devoted to the study of the cryptographic hash function MD4 and some features of its design, which allow reducing the function for the problem of restoring a preimage from a known image with some known characteristics of the preimage by the brute force method. Based on these features, a reduced parameterized MD4 algorithm is proposed, which significantly reduces the number of algorithmic and logical operations. A special case of the algorithm is the optimized preimage search algorithm by the brute-force method, proposed by the developers of the Hashcat software [1]. The generalized algorithm reduces

the number of state update steps in the inner loop of the algorithm by 52%, and also reduces the number of the most time-consuming modulo 2^{32} addition operation by 66% at best.

***Keywords*:** hash function MD4; restoring the preimage; reversible transforms; brute force method; generalized parameterized algorithm; optimization.

Введение

Криптографическая хеш-функция является одним из наиболее важных и применимых в прикладных системах криптографическим примитивом. В настоящее время наиболее используемыми в прикладных задачах хеш-функциями являются алгоритмы семейств MD и SHA. Такие алгоритмы строятся на базе парадигмы Меркла-Дамгарда [2] и повсеместно применяются в прикладных системах для безопасного хранения данных и проверки целостности.

Данная работа посвящена исследованию алгоритма MD4 из семейства алгоритмов MD. В работе исследуются некоторые особенности конструкции данного алгоритма, которые позволяют редуцировать алгоритм для задачи восстановления прообраза. Основная группа работ по исследованию поиска прообраза для хеш-функции MD4 [3, 4, 5] описывает алгоритмические подходы к решению данной задачи. В отличии от этих работ, настоящее исследование описывает обобщенный параметризированный алгоритм поиска прообраза опираясь на метод полного опробования кандидатов и предлагая технику редуцирования алгоритма для группы прообразов с известными свойствами. Один из частных случаев такой техники был описан в работе [1] разработчиками специализированного программного обеспечения для восстановления паролей Hashcat.

В настоящее время алгоритм MD4 используется в качестве криптографической хеш-функции для следующих прикладных систем:

- система хранения паролей в операционных системах типа Windows поколения NT и старше;
- система создания одноразовых паролей S/KEY [6];
- система синхронизации файлов и каталогов Rsync;
- система одноранговой сети eDonkey.

Цель настоящей работы – исследование криптографической хеш-функции MD4 и разработка алгоритма, дающего прирост скорости опробования кандидатов для метода полного опробования в задаче поиска прообраза. В ходе работы были поставлены и решены следующие задачи:

- исследование алгоритма MD4 и особенностей его конструкции;

– исследование актуальных методов и техник реализации метода полного опробования для восстановления прообраза алгоритма MD4.
– исследование возможности создания параметризированного редуцированного алгоритма для реализации метода полного опробования для восстановления прообраза алгоритма MD4.
– разработка и описание обобщенного параметризированного редуцированного алгоритма для реализации метода полного опробования для восстановления прообраза алгоритма MD4.

1. Теоретические основы

Алгоритм MD4 был предложен в качестве криптографической хеш-функции Р.Л. Ривестом и описан в работе [7] в 1990 г. Данная функция представляет из себя отображение

$$F: \{0,1\}^{(*)} \rightarrow \{0,1\}^{(128)}. \#(1)$$

Открытый текст (прообраз) M расширяется и разбивается на p блоков по k бит, а вычисление образа h происходит путем применения рекуррентной функции C (функция сжатия) над очередным промежуточным значением образа h_i и блоком открытого текста M_i:

$$C: \{0,1\}^{(128)} \times \{0,1\}^{(k)} \rightarrow \{0,1\}^{(128)}, k = 512;$$

$$h_{i+1} = C(h_i, M_i);$$

$$F(M_0, \dots, M_{p-1}) = h_p, i = \{0, \dots, p-1\}. \#(2)$$

В работе приводится описание вектора внутреннего состояния Q_j, вычисляемого на -м шаге функции C, где $Q_{-4}, Q_{-3}, Q_{-2}, Q_{-1}$ – начальные значения состояний, W_r – логическая функция над 32-х битными векторами, вектор m – 32-х битный вектор расширенного блока открытого текста и вектор t – раундовый константное 32-х битное значение. Шаги обновления состояния задаются следующими выражениями:

$$Q_0 = \left(Q_{j-4} \boxplus W_0(Q_{j-3}, Q_{j-2}, Q_{j-1}) \boxplus m_0 \boxplus t_0\right) \lll s_0, \#(3)$$

$$Q_j = \left(Q_{j-4} \boxplus W_r(Q_{j-1}, Q_{j-2}, Q_{j-3}) \boxplus m_{v_r(j)} \boxplus t_r\right) \lll s_j, j \in \{1, \dots, 47\}, \#(4)$$

Последними преобразованиями функции C являются четыре операции сложение по модулю 2^{32} векторов начального заполнения $Q_{-4}, Q_{-3}, Q_{-2}, Q_{-1}$ с векторами состояния $Q_{44}, Q_{45}, Q_{46}, Q_{47}$:

$$h_{i+1}^{(0)} = Q_{-4} \boxplus Q_{44},$$
$$h_{i+1}^{(1)} = Q_{-3} \boxplus Q_{45},$$
$$h_{i+1}^{(2)} = Q_{-2} \boxplus Q_{46},$$
$$h_{i+1}^{(3)} = Q_{-1} \boxplus Q_{47}. \#(5)$$

Для блока M_0 используются константные значения начального заполнения. Для последующих блоков M_{i+1} используются следующие начальные значения:

$$Q_{-4} = h_i^{(0)},$$
$$Q_{-3} = h_i^{(1)},$$
$$Q_{-2} = h_i^{(2)},$$
$$Q_{-1} = h_i^{(3)}. \#(6)$$

Результатом работы хеш-функции (образом) является конкатенация 32-х битных векторов:

$$h = h_p = h_{p-1}^{(0)} \,||\, h_{p-1}^{(1)} \,||\, h_{p-1}^{(2)} \,||\, h_{p-1}^{(3)}. \#(7)$$

Количество операций, используемых в стандартной реализации алгоритма MD4 приведено в Таблице 1.

Таблица 1 – Количество базовых операций в стандартной реализации алгоритма MD4

Операции над 32-х битными векторами	$\boxplus$	$\lll$	$\wedge$	$\vee$	$\oplus$	$\neg$
Количество операций	148	48	80	48	32	16

Известно, что равенство (4) является обратимым преобразованием. Обратное преобразование имеет следующий вид:

$$Q_{j-4} = (Q_j \ggg s_i) \boxminus W_r(Q_{j-1}, Q_{j-2}, Q_{j-3}) \boxminus m_{v_r(j)}^r, j \in \{0, \dots, 47\}. \#(8)$$

Кроме этого, операции сложения конечного состояния с исходным также обратимы:

$$Q_{44} = h^0 \boxminus Q_{-4},$$
$$Q_{45} = h^1 \boxminus Q_{-3},$$
$$Q_{46} = h^2 \boxminus Q_{-2},$$
$$Q_{47} = h^3 \boxminus Q_{-1}. \#(9)$$

Таким образом, зафиксировав блоки предполагаемого прообраза $m_1, ..., m_{15}$, можно получить промежуточные значения состояний из известного прообраза h:

$$Q_{47} = h^{(3)} \boxminus Q_{-1},$$
$$Q_{46} = h^{(2)} \boxminus Q_{-2},$$
$$Q_{45} = h^{(1)} \boxminus Q_{-3},$$
$$Q_{44} = h^{(0)} \boxminus Q_{-4},$$
$$Q_{43} = (Q_{47} \ggg 15) \boxminus W_2(Q_{46}, Q_{45}, Q_{44}) \boxminus m_{15}^{(2)},$$
$$\dots,$$
$$Q_{32} = (Q_{36} \ggg 3) \boxminus W_2(Q_{35}, Q_{34}, Q_{33}) \boxminus m_2^{(2)},$$
$$Q_{31} = (Q_{35} \ggg 15) \boxminus W_2(Q_{34}, Q_{33}, Q_{32}) \boxminus m_{12}^{(2)},$$
$$Q_{30} = (Q_{34} \ggg 11) \boxminus W_2(Q_{33}, Q_{32}, Q_{31}) \boxminus m_4^{(2)},$$
$$Q_{29} = (Q_{33} \ggg 9) \boxminus W_2(Q_{32}, Q_{31}, Q_{30}) \boxminus m_8^{(2)}. \#(10)$$

Обратное вычисление промежуточных значений состояний $Q_{29}, Q_{30}, Q_{31}, Q_{32}$ позволяет редуцировать алгоритм вычисления хеш-функции MD4 до вычисления только значений состояний $Q'_{29}, Q'_{30}, Q'_{31}, Q'_{32}$, сокращая количество базовых операций:

$$Q'_{-4} = 0x67452301,$$
$$Q'_{-3} = 0xefcdab89,$$
$$Q'_{-2} = 0x98badcfe,$$
$$Q'_{-1} = 0x10325476$$
$$Q'_0 = Q'_{-4} \boxplus W_0(Q'_{-3}, Q'_{-2}, Q'_{-1}) \boxplus m_0^{(0)}) \lll 3,$$
$$\dots,$$
$$Q'_{29} = (Q'_{25} \boxplus W_1(Q'_{28}, Q'_{27}, Q'_{26}) \boxplus m_7^{(1)}) \lll 5,$$
$$Q'_{30} = (Q'_{26} \boxplus W_1(Q'_{29}, Q'_{28}, Q'_{27}) \boxplus m_{11}^{(1)}) \lll 9,$$
$$Q'_{31} = (Q'_{27} \boxplus W_1(Q'_{30}, Q'_{29}, Q'_{28}) \boxplus m_{15}^{(1)}) \lll 13,$$
$$Q'_{32} = \left(Q'_{28} \boxplus W_2(Q'_{31}, Q'_{30}, Q'_{29}) \boxplus m_0^{(2)}\right) \lll 3. \#(11)$$

Используя технику обратных преобразований, разработчики специализированного программного обеспечения Hashcat, в работе [1], предлагают оптимизированный алгоритм поиска прообраза по известному образу для хеш-функции MD4 с использованием устройств ускорения вычислений. В ходе исследования было выяснено, что предложенный алгоритм можно обобщить и параметризировать для разных размеров предполагаемых прообразов.

Следует отметить, что для ускорения вычислений в задаче восстановления прообраза по известному образу используются специализированные устройства: графические ускорители, программируемые логические интегральные схемы (ПЛИС), интегральные схемы специального назначения. Алгоритм для восстановления прообраза по известному образу с использованием ускорителей вычислений реализует внешний цикл работы (на устройстве управления) и внутренний цикл работы (на устройстве ускорения вычислений). Внешний цикл работы обновляет некоторые фиксируемые значения из пространства предполагаемых векторов прообраза, а внутренний цикл производит параллельные вычисления для всего подпространства нефиксированных элементов.

2. Результаты исследования

В ходе исследования был разработан и предложен обобщенный параметризированный алгоритм восстановления прообраза по известному образу методом полного опробования кандидатов, частным случаем которого является оптимизированный алгоритм восстановления прообраза для хеш-функции MD4. Алгоритм использует особенности конструкции хеш-функций семейства MD и позволяет сократить количество операций для восстановления прообраза, тем самым значительно ускоряя процесс опробования кандидатов. Данный алгоритм является корректным за счет использования обратимых преобразований, допустимых в конструкции алгоритма хеш-функции MD4. Изменяемым входным параметром предложенного алгоритма восстановления прообраза является диапазон размера предполагаемого прообраза. Обобщенный алгоритм включает в себя 4 различных реализации в зависимости от размера l предполагаемого прообраза:

- $0 < l \leq 160$ бит;
- $160 < l \leq 288$ бит;
- $288 < l \leq 416$ бит;
- $416 < l \leq 440$ бит.

Для параметра $0 < l \leq 160$ предложенный алгоритм совпадает с алгоритмом, приведенным в работе [1].

Для описания обобщенного алгоритма для следующих параметров необходимо ввести дополнительные понятия. Пусть равенство

$$\bar{X} = \bar{X}^{(1)} \cup \ldots \cup \bar{X}^{(L)} \#(12)$$

задает объединение групп предполагаемых кандидатов-прообразов, сформированных по критерию размера кандидата, где L является максимальным размером, и выполняются условия

$$\bar{x}^{(l)} \in \bar{X}^{(l)}, l \in \{1, \ldots, L\}, \#(13)$$

где

$$\bar{x}^{(l)} = \left(x_0^{(l)}, \ldots, x_{l-1}^{(l)}\right), x_q^{(l)} \in \{0,1\}. \#(14)$$

Равенство

$$\bar{X}^{(l)} = \bar{Z}^{(n)} \,||\, \bar{V}^{(l-n)}, n < l \#(15)$$

определяет конкатенацию частей кандидатов-прообразов для внутреннего цикла $\bar{Z}^{(n)}$ и внешнего цикла $\bar{V}^{(l-n)}$ вычислений.

Для параметра $160 < l \leq 288$, зафиксировав во внешнем цикле все блоки сообщения, кроме блока m_4, появляется возможность предварительно вычислить состояния Q_0, Q_1, Q_2, Q_3, а также путем обратного преобразования вычислить состояния $Q_{31}, Q_{32}, Q_{33}, Q_{34}$. Кроме этого, для корректного описания алгоритма необходимо определить следующие равенства:

$$\bar{X}^{(l)} = \bar{V}_0^{(128)} \,||\, \bar{Z}^{(n)} \,||\, \bar{V}_1^{(l-n-128)}, \#(16)$$

$$\bar{V}^{(l-n)} = \bar{V}_0^{(128)} \,||\, \bar{V}_1^{(l-n-128)}. \#(17)$$

Алгоритм – Восстановление прообраза MD4 по известному образу с использованием ускорителей вычислений для параметра $160 < l \leq 288$

Вход: $h, i, l, \bar{X}$.
Выход: $\bar{x}$ или $\emptyset$.
Алгоритм:

1. Пока $l \leq L$:
   1.1 Для $\bar{X}^{(l)}$ зафиксировать $n, n \leq 32$.
   1.2 Для каждого $\bar{v}_q^{(l-n)} \in \bar{V}^{(l-n)}$ ***(внешний цикл)***:
      1.2.1 Зафиксировать $m_0, \ldots, m_{15}$; $m_4 = 0$.
      1.2.2 Вычислить $m^{(0)}, m^{(1)}, m^{(2)}$.
      1.2.3 Вычислить Q_0, Q_1, Q_2, Q_3.

1.2.4 Вычислить $Q_{31}, Q_{32}, Q_{33}, Q_{34}$.

1.2.5 Для каждого $\bar{z}_p^{(n)} \in \bar{Z}^{(n)}$ ***(внутренний цикл, параллельно)***:

1.2.5.1 Зафиксировать m_4.

1.2.5.2 Вычислить $m_4^{(0)}, m_4^{(1)}, m_4^{(2)}$.

1.2.5.3 Для состояний Q_0, Q_1, Q_2, Q_3 вычислить $Q_{31}^{'}, Q_{32}^{'}, Q_{33}^{'}, Q_{34}^{'}$.

1.2.5.4 Если выполняются равенства

$$Q'_{34} = Q_{34},$$
$$Q'_{33} = Q_{33},$$
$$Q'_{32} = Q_{32},$$
$$Q'_{31} = Q_{31},$$

вернуть $\bar{x}^{(l)}$ и закончить работу.

1.3 Увеличить l.

2. Вернуть $\emptyset$ и закончить работу.

Результатом использования данного алгоритма является сокращение количества шагов обновления состояния во внутреннем цикле до 30 – на 38% меньше в сравнении со стандартной реализацией алгоритма хеш-функции MD4.

Для параметра $288 < l \leq 416$ справедливы следующие равенства:

$$\bar{X}^{(l)} = \bar{V}_0^{(256)} \,||\, \bar{Z}^{(n)} \,||\, \bar{V}_1^{(l-n-256)}, \#(18)$$

$$\bar{V}^{(l-n)} = \bar{V}_0^{(256)} \,||\, \bar{V}_1^{(l-n-256)}. \#(19)$$

Зафиксировав во внешнем цикле все блоки сообщения, кроме блока m_8, появляется возможность предварительно вычислить состояния Q_4, Q_5, Q_6, Q_7, а также путем обратного преобразования вычислить состояния $Q_{30}, Q_{31}, Q_{32}, Q_{33}$.

Алгоритм – Восстановление прообраза MD4 по известному образу с использованием ускорителей вычислений для параметра $288 < l \leq 416$

Вход: $h, i, l, \bar{X}$.
Выход: $\bar{x}$ или $\emptyset$.
Алгоритм:

1. Пока $l \leq L$:

1.1 Для $\bar{X}^{(l)}$ зафиксировать $n, n \leq 32$.

1.2 Для каждого $\bar{v}_q^{(l-n)} \in \bar{V}^{(l-n)}$ ***(внешний цикл)***:

1.2.1 Зафиксировать $m_0, \ldots, m_{15}; m_8 = 0$.

1.2.2 Вычислить $m^{(0)}, m^{(1)}, m^{(2)}$.

1.2.3 Вычислить Q_4, Q_5, Q_6, Q_7.

1.2.4 Вычислить $Q_{30}, Q_{31}, Q_{32}, Q_{33}$.

1.2.5 Для каждого $\bar{z}_p^{(n)} \in \bar{Z}^{(n)}$ ***(внутренний цикл, параллельно)***:

1.2.5.1 Зафиксировать m_8.

1.2.5.2 Вычислить $m_8^{(0)}, m_8^{(1)}, m_8^{(2)}$.

1.2.5.3 Для состояний Q_4, Q_5, Q_6, Q_7 вычислить $Q_{30}^{'}, Q_{31}^{'}, Q_{32}^{'}, Q_{33}^{'}$.

1.2.5.4 Если выполняются равенства

$$Q'_{33} = Q_{33},$$
$$Q'_{32} = Q_{32},$$
$$Q'_{31} = Q_{31},$$
$$Q'_{30} = Q_{30},$$

вернуть $\bar{x}^{(l)}$ и закончить работу.

1.3 Увеличить l.

2. Вернуть $\emptyset$ и закончить работу.

Для параметра $288 < l \leq 416$ происходит сокращение количества шагов обновления состояния во внутреннем цикле до 26 – на 46% меньше в сравнении со стандартной реализацией алгоритма хеш-функции MD4.

Наиболее эффективно алгоритм работает для параметра $416 < l \leq 440$. В данном случае справедливы следующие равенства:

$$\bar{X}^{(l)} = \bar{V}_0^{(384)} \,||\, \bar{Z}^{(n)} \,||\, \bar{V}_1^{(l-n-384)}, \#(20)$$

$$\bar{V}^{(l-n)} = \bar{V}_0^{(384)} \,||\, \bar{V}_1^{(l-n-384)}. \#(21)$$

Зафиксировав во внешнем цикле все блоки сообщения, кроме блока m_{12}, появляется возможность предварительно вычислить состояния Q_8, Q_9, Q_{10}, Q_{11}, а также путем обратного преобразования вычислить состояния $Q_{32}, Q_{33}, Q_{34}, Q_{35}$.

Алгоритм – Восстановление прообраза MD4 по известному образу с использованием ускорителей вычислений для параметра $416 < l \leq 440$

Вход: $h, i, l, \bar{X}$.
Выход: $\bar{x}$ или $\emptyset$.
Алгоритм:

1. Пока $l \leq L$:

1.1 Для $\bar{X}^{(l)}$ зафиксировать $n, n \leq 32$.

1.2 Для каждого $\bar{v}_q^{(l-n)} \in \bar{V}^{(l-n)}$ ***(внешний цикл)***:

1.2.1 Зафиксировать $m_0, ..., m_{15}$; $m_{12} = 0$.

1.2.2 Вычислить $m^{(0)}, m^{(1)}, m^{(2)}$.

1.2.3 Вычислить Q_8, Q_9, Q_{10}, Q_{11}.

1.2.4 Вычислить $Q_{32}, Q_{33}, Q_{34}, Q_{35}$.

1.2.5 Для каждого $\bar{z}_p^{(n)} \in \bar{Z}^{(n)}$ ***(внутренний цикл, параллельно)***:

1.2.5.1 Зафиксировать m_{12}.

1.2.5.2 Вычислить $m_{12}^{(0)}, m_{12}^{(1)}, m_{12}^{(2)}$.

1.2.5.3 Для состояний Q_8, Q_9, Q_{10}, Q_{11} вычислить $Q_{32}^{'}, Q_{33}^{'}, Q_{34}^{'}, Q_{35}^{'}$.

1.2.5.4 Если выполняются равенства

$$Q'_{35} = Q_{35},$$
$$Q'_{34} = Q_{34},$$
$$Q'_{33} = Q_{33},$$
$$Q'_{32} = Q_{32},$$

вернуть $\bar{x}^{(l)}$ и закончить работу.

1.3 Увеличить l.

2. Вернуть $\emptyset$ и закончить работу.

Для параметра $416 < l \leq 440$ происходит сокращение количества шагов обновления состояния во внутреннем цикле до 23, что на 52% меньше в сравнении со стандартной реализацией.

Корректность алгоритмов данного типа определяется свойством обратимости преобразований алгоритма хеш-функции MD4. Оценка трудоемкости алгоритмов зависит от выбранных параметров и опирается на оценку трудоемкости стандартного алгоритма поиска прообраза по известному образу с использованием ускорителей вычислений $O(\sum_{l=1}^{m} | \bar{V}_{l-n} |)$.

Заключение

В Таблице 2 приводится сравнение наиболее значимых характеристик обобщенного параметризированного алгоритма с различными параметрами.

Предложенный алгоритм позволяет сократить количество шагов обновления состояния во внутреннем цикле до 52% в сравнении со стандартной реализацией, тем самым сокращая количество базовых операций, в том числе наиболее трудоемкой операции сложения по модулю 2^{32} до 66%.

Таблица 2 – Некоторые характеристики обобщенного алгоритма для различных параметров

Параметр	Количество шагов обновления состояния во внутреннем цикле	Сокращение количества шагов обновления состояния во внутреннем цикле в сравнении со стандартной реализацией	Количество операций сложения по модулю 2^{32} ($\boxplus$) во внутреннем цикле
$0 < l \leq 160$	33	31%	68
$160 < l \leq 288$	30	38%	64
$288 < l \leq 416$	26	46%	54
$416 < l \leq 440$	23	52%	50

Построение алгоритмов подобного типа возможно для некоторых других хеш-функций семейства MD и SHA в связи с обратимостью преобразования обновления состояния и схожести конструкции данных хеш-функций с исследуемой хеш-функцией MD4.

Библиографические ссылки

1. Steube J. Optimizing computation of Hash-Algorithms as an attacker. 2013. URL: https://hashcat.net/events/p13/js-ocohaaaa.pdf.
2. Lai X., Massey J.L.: Hash Function Based on Block Ciphers // EUROCRYPT. 1992. С. 55–70.
3. Dobbertin H.: The first two rounds of MD4 are not one-way. In: Vaudenay, S. (ed.) FSE 1998. LNCS, vol. 1372, pp. 284–292. Springer, Heidelberg (1998).
4. De D., Kumarasubramanian A., Venkatesan R.: Inversion attacks on secure hash functions using sat solvers. In: Marques-Silva, J., Sakallah, K.A. (eds.) SAT 2007. LNCS, vol. 4501, pp. 377–382. Springer, Heidelberg (2007).
5. Kuwakado H., Tanaka H.: New algorithm for finding preimages in a reduced version of the MD4 compression function // IEICE Transactions Fundamentals of Electronics, Communications and Computer Sciences (Japan). 2000. № E83-A(1). P. 97–100.
6. Haller N. The S/KEY One-Time Password System // RFC 1760 (Informational) (February 1995).
7. Rivest R.L. The MD4 Message Digest Algorithm. In Menezes A., Vanstone S.A., eds.: CRYPTO. Vol. 537 of Lecture Notes in Computer Science. Springer. 1990. P. 303–311.

УДК 004.056

СТРУКТУРА И СОСТАВ КОРПОРАТИВНЫХ ПОДРАЗДЕЛЕНИЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В.П. Кочин[1], А.В. Шанцов[2]

[1]Белорусский государственный университет,
пр-т Независимости, 4, 220030, г. Минск, Беларусь, Kochyn@bsu.by
[2]Белорусский государственный университет,
пр-т Независимости, 4, 220030, г. Минск, Беларусь, ShantsovAV@bsu.by

Рассмотрена проблематика обеспечения безопасности информационных ресурсов в Республики Беларусь. Определена необходимость применения подразделений информационной безопасности для обеспечения защиты информационных ресурсов. Рассмотрены типы подразделений информационной безопасности. Определены цели и задачи корпоративного подразделения информационной безопасности. Рассмотрены требования к инфраструктуре и средствам автоматизации подразделения информационной безопасности корпоративного уровня. Определен основной состав корпоративного подразделения информационной безопасности. Предложены структуры корпоративных подразделений информационной безопасности. Приведены краткие характеристики подразделений информационной безопасности корпоративного уровня в зависимости от масштабов защищаемых информационных ресурсов.

Ключевые слова: информационные технологии; информационная безопасность; подразделение информационной безопасности.

STRUCTURE AND COMPOSITION OF CORPORATE INFORMATION SECURITY UNITS

V.P. Kochyn[a], A.V. Shantsou[b]

[a]Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus,
Kochyn@bsu.by
[b]Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus,
ShantsovAV@bsu.by

The problems of ensuring the security of information resources in the Republic of Belarus are considered. The necessity of using information security units to ensure the protection of information resources is determined. The types of information security units are considered. The goals and objectives of the corporate information security unit are defined. The requirements for the infrastructure and automation tools of the corporate-level information security unit are considered. The main composition of the corporate information security unit has been determined. The structures of corporate information security unit are proposed. Brief characteristics of corporate-level information security units are given depending on the scale of protected information resources.

Keywords: information technology; information security; information security unit.

Введение

В настоящее время информационные ресурсы (далее – ИР) подвергаются постоянным угрозам информационной безопасности (далее – ИБ). Применение хорошо продуманных архитектур защиты информации не позволяет гарантировать полную защищенность ИР [1]. Новые уязвимости в программном и аппаратном обеспечении, позволяют злоумышленникам разрабатывать новые методы преодоления средств защиты. На сегодняшний день в Республике Беларусь технические нормативные правовые акты по защите информации предъявляют требования по обеспечению ИБ лишь пассивными методами [2,3]. Однако, для полноценной защиты ИР пассивных методов защиты недостаточно, необходимо постоянно отслеживать актуальное состояние защищенности ИР, регулярно внедрять новые методы и механизмы защиты.

Вышеперечисленные факторы приводят к необходимости внедрения комплексного подхода по защите ИР [4]. Данный подход в том числе предполагает защиту ИР с помощью подразделений ИБ, способных обнаруживать инциденты ИБ, реагировать на них, проводить расследования данных инцидентов, участвовать в ликвидации их последствий, проводит постоянный мониторинг событий ИБ, внедрять новые методы и механизмы защиты информации. В мировой практике, в зависимости от решаемых задач и масштабов защищаемых информационных ресурсов, данные подразделения именуется как Security Operation Center (SOC), Computer Emergency Response Team (CERT) или Computer Security Incident Response Team (CSIRT). Назначение и задачи подразделений ИБ существенно отличаются в зависимости от типа подразделения ИБ. Выделяют несколько типов подразделений ИБ:

подразделения ИБ национального уровня;

иерархические (отраслевые) подразделения ИБ;

корпоративные подразделения ИБ.

Актуальность проблемы защиты ИР с помощью подразделений ИБ обусловлено также тем, что законодательство Республики Беларусь в ряде случаев требует от владельцев ИР иметь в своем составе подразделение, отвечающее за ИБ [5], однако цели и задачи подразделений ИБ, а также методика расчета структуры подразделений ИБ не определены.

Методология исследования

В настоящей статье анализу подвергается подразделение ИБ корпоративного уровня, непосредственно обеспечивающее защиту ИР корпора-

ции (организации) (далее – подразделение ИБ). Анализ основывается на теоретических основах по защите информации, методикам противодействия кибератакам, а также лучшим мировым практикам по созданию подразделений ИБ.

Цели и задачи подразделения ИБ

Целью подразделения ИБ является обеспечение информационной безопасности защищаемых ИР. Для достижения данной цели, подразделение ИБ решает задачи в зависимости от структуры и масштабов защищаемых ИР, а также необходимого уровня защиты ИР. Основными задачами, решаемыми подразделением ИБ, являются:

инвентаризация защищаемых активов и поддержание информации о них в актуальном состоянии;

мониторинг и обработка событий и сетевого трафика в режиме реального времени;

анализ действий пользователей;

анализ уязвимостей и контроль их устранения;

поведенческий анализ файлов;

проверка устойчивости информационных ресурсов к внешним воздействиям (тесты на проникновение);

настройка и управление датчиками аналитических систем, техническое обслуживание инфраструктуры подразделения ИБ;

обработка информации об инцидентах ИБ, проведение расследований инцидентов ИБ, анализ инцидентов ИБ и реагирование на них;

проведение тренировок по реагированию на инциденты ИБ;

разработка и внедрение инструментов подразделения ИБ.

Подразделение ИБ, в зависимости от защищаемых ИР и требований по их защите, решает либо часть из перечисленных задач, либо решает задачи в полном объеме, а также в отдельных случаях решает дополнительные задачи для поддержания требуемого уровня защиты ИР.

Для решения поставленных задач, подразделения ИБ должно включать: инфраструктуру, средства автоматизации и персонал.

Инфраструктура и средства автоматизации подразделения ИБ

При функционировании подразделения ИБ немаловажную роль играет инфраструктура. Необходимо предусмотреть масштабирование инфраструктуры подразделения ИБ для наращивания объемов хранения информации, усиления отказоустойчивости при сборе событий ИБ и

увеличения скорости обработки данных. Обязательным является подготовка инфраструктуры к хранению больших объемов данных.

К средствам автоматизации подразделений ИБ относят такие системы, как система сбора и анализа информации о событиях информационной безопасности SIEM (Security Information and Event Management), система оркестрировки, автоматизации и реагирования SOAR (Security Orchestration Automation and Response), система сбора событий LM (Log Management). Применение систем SIEM и SOAR является приоритетной задачей, так как позволяет существенно сократить затраты на человеческие ресурсы по сравнению с другими средствами автоматизации.

Дополнительно в подразделении ИБ используются такие средства автоматизации, как средства реверс-инжиниринга, средства сбора криминалистических доказательств, сетевые сканеры, сканеры уязвимостей и другие. Применение средств автоматизации в составе подразделения ИБ позволяет существенно снизить штат подразделения ИБ, что в свою очередь позволяет инвестировать в подготовку персонала вместо найма большего количества. Модель подразделения ИБ с меньшим числом подготовленных специалистов имеет экономические преимущества по сравнению с моделью с большим числом сотрудников среднего уровня подготовленности [6].

Персонал подразделения ИБ

В независимости от структуры подразделения ИБ, для эффективного противодействия атакам на защищаемые ИР, подразделение ИБ должно осуществлять реагирование на инциденты ИБ в масштабе времени, соответствующим масштабу времени действий нарушителя [7, с. 41]. Для соблюдения данного критерия, в одной организационной структуре подразделения ИБ должно обеспечиваться выполнение следующих задач:

мониторинг и обработка событий в режиме реального времени;

анализ инцидентов ИБ и реагирование на них;

настройка и управление датчиками аналитических систем, техническое обслуживание инфраструктуры подразделения ИБ.

Для решения перечисленных задач в составе подразделения ИБ должны быть:

- Руководитель подразделения ИБ. Руководитель подразделения ИБ осуществляет общее руководство работой подразделения ИБ, координирует взаимодействие между командой подразделения ИБ. Осуществляет подбор кадров, управление финансами подразделения ИБ, проводит работу с потенциальными и существующими клиентами.

- Аналитик 1-го уровня. Аналитик 1-го уровня осуществляет первичную обработку инцидентов ИБ – распределение и первичный отсев ложных срабатываний. Аналитики 1-го уровня в работе опираются на заранее созданные администратором безопасности сценарии реагирования, в которых указана последовательность шагов, которые надо предпринять при обнаружении того или иного типа инцидента ИБ. В случае, если аналитик 1-го уровня может самостоятельно выполнить все действия, он осуществляет реагирование своими силами. Если он столкнулся с необходимостью эскалации инцидента ИБ, то он передает его на следующий уровень – аналитику 2-го уровня. На аналитиков 1-го уровня также возлагается взаимодействие с клиентами (прием заявок, звонков).
- Аналитик 2-го уровня. Аналитик 2-го уровня получает данные об инцидентах от аналитика 1-го уровня. Аналитик 2-го уровня не опирается на какие-либо шаблоны реагирования, а анализирует уникальную ситуацию, сопоставляя различные события и факты, имеющие отношение к инциденту ИБ. В случае, если в расследуемой кибератаке применялось ранее неизвестное вредоносное ПО или применен новый тип атаки, аналитик 2-го уровня, при реагировании на инцидент ИБ, взаимодействует с реверс-инженерами и экспертами-криминалистами (при их наличии в составе подразделения ИБ).
- Администратор безопасности. Администратор безопасности отвечает за настройку правил в системах SIEM (SOAR). Администратор безопасности получает исходные данные о работе информационных систем и составляет правила выявления инцидентов и реагирования на них. Это правила корреляции в системах SIEM, которые отвечают за обработку входящих сообщений от систем безопасности и за выстраивание этих разнородных событий в логически целостную «историю» для поиска возможной атаки, а также правила автоматического реагирования, локализации, восстановления информационных систем при помощи SOAR решений. Совместно с аналитиками 2-го уровня осуществляет написания сценариев реагирования для аналитиков 1-го уровня. В случае, если для защиты используются системы обнаружения/предотвращения вторжений, то на администратора безопасности возлагаются функции по написанию для них сигнатур.
- Системный администратор (системный инженер) – лицо ответственное за настройку и поддержание работоспособности инфраструктуры подразделения ИБ. В обязанности системного администратора входит конфигурирование различных типов операционных систем, прикладного программного обеспечения, систем

защиты, сетевого оборудования. В случае, если в подразделении ИБ используется какое-то самостоятельно созданное программное обеспечение (далее – ПО), то на системного администратора (системного инженера) возлагают функции непрерывной интеграции и настройки такого ПО (функции DevOps).

Наличие указанных должностей в составе подразделения ИБ является обязательным. Важно отметить, что выше приведены именно должности, а не состав подразделения ИБ. Для повышения уровня защищенности ИР и повышения уровня обслуживания в составе подразделения ИБ дополнительно предусматриваются следующие должности:

- Специалист по реверс-инжинирингу (реверс-инженер). Реверс-инженер выполняет задачи по изучению образцов вредоносного ПО для противодействия осуществляемым с их помощью кибератакам. При выявлении нового вредоносного ПО, на реверс-инженера возлагаются функции по написанию сигнатур вредоносного ПО.
- Эксперт-криминалист. На эксперта-криминалиста возлагаются задачи по поиску изменений (аномалий) в атакованной системе, определение данных, которые были удалены, изменены или похищены, определение систем, которые были атакованы. На эксперта-криминалиста возлагаются задачи по воссозданию полного пути распространения атаки: что именно подвергалось атаке, как развивалась атака и какой был причинен ущерб.
- Специалист по кибербразведке. Специалист по кибербразведке отвечает за поиск уязвимостей в защищаемых ИР, по согласованию с владельцами ИР выполняет тестирование на проникновение. Дополнительно специалист по кибербразведке осуществляет аудит защищаемых ИР с целью обнаружения вредоносных программ в системах, например вирусов-логических бомб, которые срабатывают только при наступлении определенных условий, а до этого никак себя не проявляют. На данного специалиста возлагают функции по поиску информации о новых типах атак и уязвимостей.

Рассмотренный перечень должностей команды подразделения ИБ позволяет решать все задачи, возлагаемые на подразделение ИБ. В определенных случаях обязанности каждого из членов команды могут дополняться в зависимости от решаемых задач.

Структура подразделения ИБ

Подразделение ИБ может быть как независимой организацией, оказывающей услуги в форме аутсорсинга, так и структурно входить в состав

компании, ИР которой подлежат защите. В отдельных случаях, подразделение ИБ не будет выделяться как таковое, вместо этого задачи по защите ИР будут возложены на системных администраторов ИР. На структуру подразделения ИБ оказывает влияние необходимый уровень обслуживания (защиты) ИР и допустимый уровень затрат на защиту ИР.

Для эффективного функционирования и соблюдения требований клиентов, структура подразделения ИБ должна соответствовать трем критериям [7, с. 49]:

1. Подразделение ИБ должно обладать квалифицированным персоналом в области защиты информации.
2. Подразделение ИБ, в частности инфраструктура, должна обеспечивать максимальную эффективность при сборе и обработке событий ИБ, анализе циркулирующего трафика в ИР.
3. Подразделение ИБ должно учитывать экономические ограничения клиентов по защите ИР и ограничения полномочий подразделения ИБ при предотвращении атак на ИР.

Выбор верной структуры подразделения ИБ позволяет соблюсти баланс между вышеперечисленными критериями и обеспечить максимальную эффективность функционирования подразделения ИБ. Рассмотрим варианты структур подразделений ИБ в зависимости от масштабов защищаемых ИР.

Нижняя граница защищаемых ИР структурно-независимым подразделением ИБ составляет порядка 2000 – 5000 устройств (узлов). Данное подразделение ИБ является либо независимой организацией, либо структурным подразделением компании-владельца ИР. Для решения задач по защите ИР указанных масштабов подразделение ИБ должно обладать собственной инфраструктурой, средствами автоматизации (SIEM или SOAR). Следовательно, в составе подразделения ИБ будет администратор безопасности и системный администратор.

Масштаб защищаемых информационных ресурсов до 10 000 узлов, позволит эффективно функционировать, с экономической точки зрения, только в режиме 8/5 (5 дней в неделю по 8 часов). При этом, для обеспечения защиты ИР масштабов от 2 000 до 10 000 устройств (узлов), потребуется две группы из аналитиков 1-го и 2-го уровней в соотношении два аналитика 1-го уровня на одного аналитика 2-го уровня.

Предлагаемая структура подразделения ИБ представлена на рисунке 1.

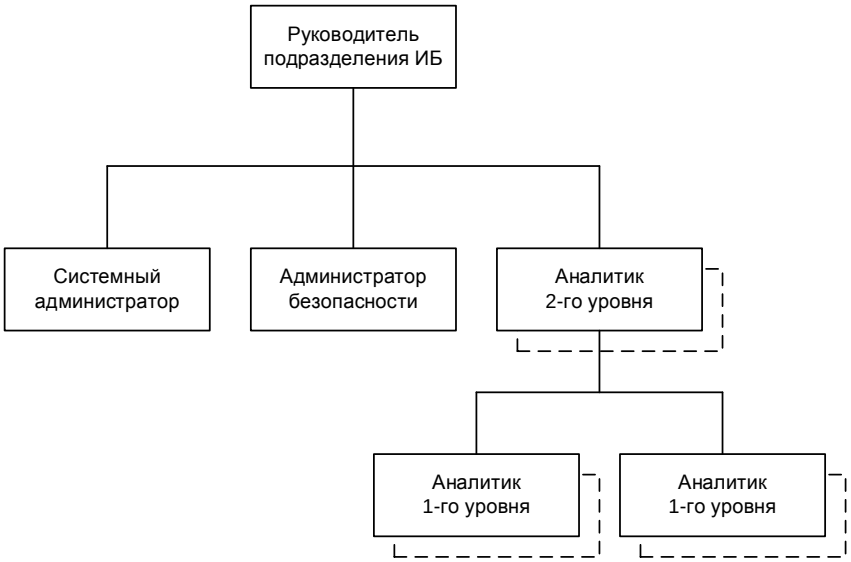


Рисунок 1 – Структура подразделения ИБ

Увеличение масштабов защищаемых ИР или необходимость в предоставлении более высокого уровня обслуживания при защите ИР приводят к необходимости расширения подразделения ИБ, а также к необходимости перехода к круглосуточному режиму работы подразделения ИБ (части подразделения ИБ).

Как правило ИР с масштабами более 10 000 устройств (узлов) не являются однородными и состоят из множества информационных систем (далее – ИС). Для повышения эффективности защиты ИС проводятся тестирования на проникновение, выявляющие слабые места в системе защиты ИС. За каждой ИС закрепляется группа из аналитиков 1-го и 2-го уровней, обладающих максимальной осведомленностью о составе и структуре, защищаемой ИС. Также, высокий уровень защищенности ИР подразумевает наличие в составе подразделения ИБ реверс-инженера и эксперта-криминалиста. Основными задачами данных специалистов является экспертная поддержка аналитиков 2-го уровня при расследовании сложных инцидентов ИБ.

Увеличение масштабов защищаемых ИР приводит к увеличению масштабов инфраструктуры подразделения ИБ, количества датчиков аналитическим систем и числа собираемых событий ИБ. Следовательно, увеличиться и количество администраторов безопасности и системных администраторов.

Круглосуточный режим работы подразделения ИБ подразумевает круглосуточную работу аналитиков 1-го уровня. При этом режим работы аналитиков 2-го уровня варьируется: 8/5 или 24/7. На практике, из-за отсутствия необходимости в столь высоком уровне обслуживания и эконо-

мической целесообразности, чаще используется режим работы аналитиков 2-го уровня 8/5. Таким образом, будут применяться группы из аналитиков 1-го и 2-го уровней в соотношении три аналитика 1-го уровня на одного аналитика 2-го уровня. При этом, под аналитиками 1-го уровня подразумевается рабочее место, так как при режиме работы 24/7 на одно рабочее место будет приходится 4 – 5 сотрудников. Количество групп аналитиков 1-го и 2-го уровней зависит от масштабов защищаемых информационных ресурсов и может варьироваться в пределах от 2 до 4 групп. В предлагаемой структуре подразделения ИБ выделено 3 таких группы.

Предлагаемая структура расширенного подразделения ИБ приведена на рисунке 2.

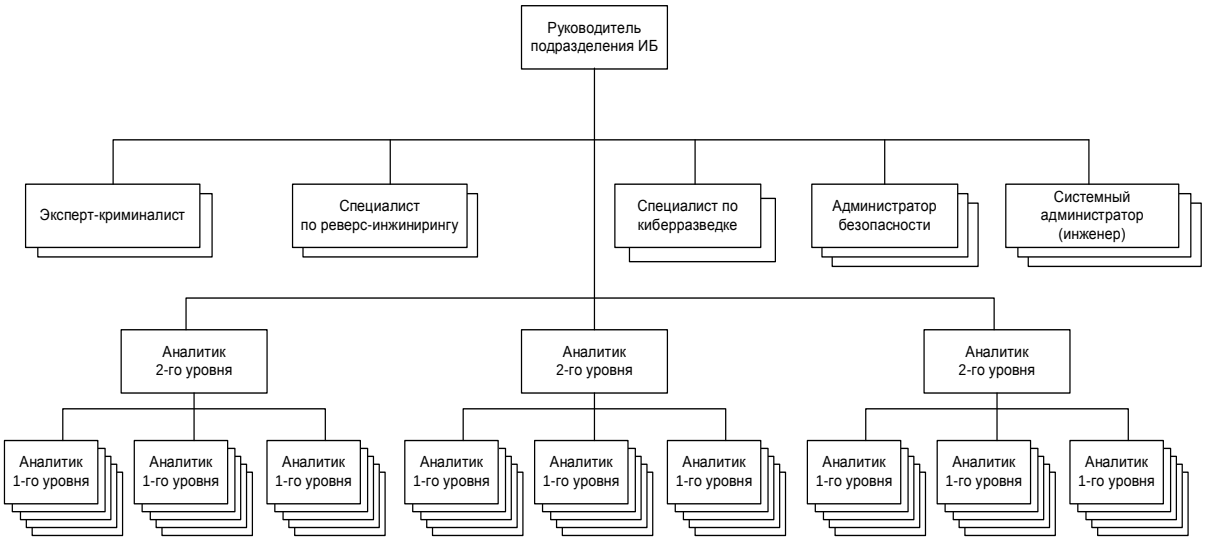


Рисунок 2 – Структура расширенного подразделения ИБ

При размерах защищаемых ИР менее 2 000 устройств (узлов) защита с использованием полноценного подразделения ИБ является экономически нецелесообразной. Для обеспечения защиты ИР данных масштабов применяют «виртуальное» подразделение ИБ. Термин «виртуальное» подразумевает, что подразделение ИБ не имеет необходимой команды в полном составе, либо команда подразделения ИБ частично состоит из специалистов совмещающих выполнение обязанностей нескольких должностей, в том числе и за пределами подразделения ИБ.

«Виртуальное» подразделение ИБ использует инфраструктуру ИР, следовательно, отсутствует необходимость в системном администраторе. Если настройка средств защиты осуществляется системными администраторами ИР, то также отсутствует необходимость в администраторе безо-

пасности. В данном случае «виртуальное» подразделение ИБ будет состоять лишь из аналитиков 1-го и 2-го уровней, при этом на аналитика 2-го уровня возлагаются задачи по руководству командой. Предлагаемая структура «виртуального» подразделения ИБ представлена на рисунке 3.

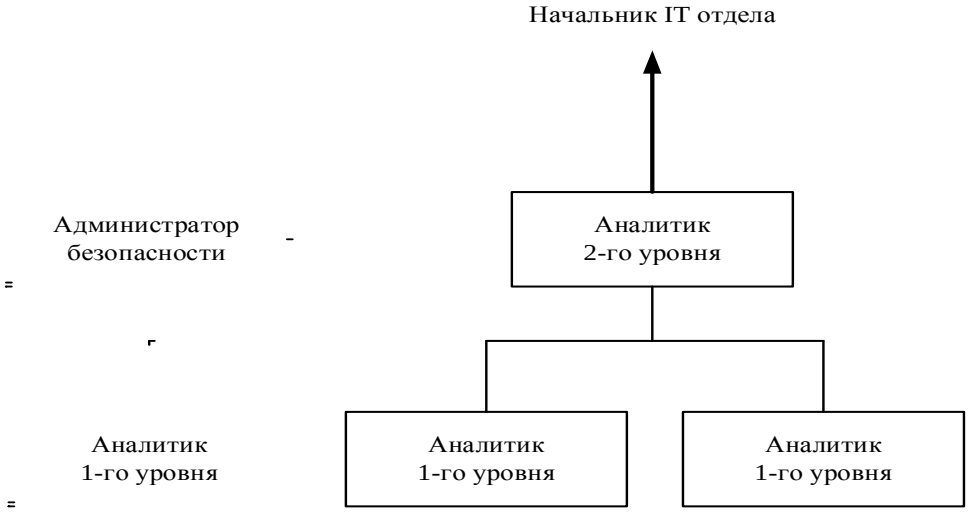


Рисунок 3 – Структура «виртуального» подразделения ИБ

При защите ИР с размерами порядка 500 устройств (узлов), применение подразделений ИБ является нецелесообразным. Данные масштабы ИР позволяют системным администраторам ИР самостоятельно поддерживать необходимый уровень защищенности. При этом необходимо, чтобы из команды системных администраторов было назначено лицо, непосредственно отвечающее за защиту ИР.

Обобщенная характеристика подразделений ИБ в зависимости от масштабов защищаемых ИР приведена в таблице.

Таблица – Характеристики подразделений ИБ

Подразделение ИБ	Ориентировочные размеры защищаемых ИР	Краткая характеристика
Защита системными администраторами ИР	До 500	Обеспечение защиты ИР командой системных администраторов; назначение из состава системных администраторов ответственного лица за защиту ИР.
«Виртуальное» подразделение ИБ	500 – 3 000	Отсутствие или совмещение ряда должностей в составе подразделения ИБ; отсутствие собственной инфраструктуры, использование инфраструктуры защищаемых ИР; отсутствие аналитических систем (SIEM, SOAR); отсутствие полномочий по самостоятельному

		реагированию на выявленные инциденты ИБ; входит в состав компании владельца информационных ресурсов.
Подразделение ИБ	3 000 – 10 000	Наличие собственной инфраструктуры; применение аналитических систем сбора и анализа событий (SIEM, SOAR); совместное реагирование на инциденты вместе с владельцами ИР; функционирование в режиме 8/5; входит в состав компании владельца ИР или является самостоятельной организацией.
Расширенное подразделение ИБ	От 10 000	Наличие собственной инфраструктуры; применяются системы автоматизации сбора и анализа событий, инструменты киберкриминалистики, реверс-инжиниринга и другие, в том числе самостоятельно разработанные; обладает полномочиями по самостоятельному реагированию на инциденты ИБ; частично функционирует в режиме 24/7; обслуживает крупные компании, информационные ресурсы которых часто подвергаются атакам и требуют высокого уровня защиты.

Заключение

В настоящем анализе определен состав и предложены структуры подразделений ИБ. Рассмотрена зависимость структуры подразделения ИБ от масштабов защищаемых ИР, необходимого уровня обслуживания и максимально возможного уровня затрат. Предложены структуры подразделений ИБ в зависимости от масштабов защищаемых ИР и решаемых задач.

Приведенные границы масштабов ИР для выбора типа подразделения ИБ не являются точными и носят рекомендательный характер. Например, ИР государственных учреждений, ИР, обрабатывающие персональные данные, априори требуют более высокий уровень защиты независимо от их масштабов. При определении того или иного типа подразделения ИБ, владельцы ИР должны определить, что для них является более приоритетным: высокий уровень защищенности при высоком уровне затрат, или достаточный уровень защищенности при среднем уровне затрат.

Активная защита информации с помощью подразделений ИБ играет важную роль при защите ИР. Непрерывное совершенствование методов проведения атак на ИР, создание и развитие новых средств для проведения кибератак делают уязвимыми пассивные системы защиты информа-

ции. Для обеспечения комплексности, система защиты информации должна состоять как из пассивной, так и из активной составляющих. Таким образом, активная защита с помощью подразделений ИБ является неотъемлемой частью комплексной системы защиты информации.

Библиографические ссылки

1. Кочин В.П., Шанцов А.В., Проблемы проектирования комплексной системы защиты информации облачных ресурсов в Республике Беларусь // Цифровая трансформация. 2021; № 3. С. 34–39.
2. Приказ Оперативно-аналитического центра при Президенте Республики Беларусь от 20 февраля 2020 г. № 66 «О мерах по реализации Указа Президента Республики Беларусь от 9 декабря 2019 г. № 449». Минск: Оперативно-аналитический центр при Президенте Республики Беларусь, 2021.
3. Приказ Оперативно-аналитического центра при Президенте Республики Беларусь от 12 ноября 2021 г. № 195 «О технической и криптографической защите персональных данных». Минск: Оперативно-аналитический центр при Президенте Республики Беларусь, 2021.
4. Кочин В.П., Шанцов А.В., Комплексная система защиты информации облачных ресурсов // Материалы XXVI научно-практической конференции «Комплексная защита информации». НП РУП «Научно-исследовательский институт технической защиты информации». 2021. С. 332 – 334.
5. Закон Республики Беларусь от 7 мая 2021 г. № 99-З «О защите персональных данных». Минск: Национальный правовой Интернет-портал Республики Беларусь, 2021.
6. Rajnovic D. Computer Incident Response and Product Security. Indianapolis: Cisco Press, 2011.
7. Zimmerman C. Ten strategies of a world-class cybersecurity operations center. Bedford: MITRE, 2014.

УДК 004.428.4

РАЗРАБОТКА ПРОГРАММНОЙ СИСТЕМЫ КОНТРОЛЯ ФОРМЫ ПЕРИМЕТРА ПОКРЫТИЯ БЕСПРОВОДНОЙ СЕТИ

В.Н. Кулинченко, Д.Ю. Путьков

Гомельский государственный университет имени Франциска Скорины, ул. Советская, 104, 246028, г. Гомель, Республика Беларусь, kulinchenko@gsu.by

В статье описывается создание веб-приложения по контролю формы периметра покрытия беспроводной сети Wi-Fi на поэтажном плане здания. Исследования зон покрытия и мощности информационных сигналов проводились для сетевых устройств стандарта 802.11 с заданием их параметров пользователем. Исследование предполагает разработку графического 2D отображения зон Wi-Fi сигнала на этажах здания, что в конечном итоге позволит правильно геометрически визуализировать зоны покрытия точек и определить выход их сигналов за периметр здания, что является потенциально опасным с точки зрения безопасности сети Wi-Fi.

Ключевые слова. веб-приложение**;** клиент; сервер; сеть Wi-Fi; безопасность; дальность действия точки доступа; препятствия.

DEVELOPMENT OF A SOFTWARE SYSTEM FOR CONTROLLING THE SHAPE OF THE WIRELESS NETWORK COVERAGE PERIMETER

V.N. Kulinchenko, D.Y. Putkov

Francisk Skorina Gomel State University, 104 Sovetskaya str., 246019, Gomel, Belarus, kulinchenko@gsu.by

This article describes the creation of a web application to control the shape of the Wi-Fi coverage perimeter on the floor plan of the building. Studies of coverage areas and information signal strength were conducted for 802.11 network devices with the setting of their parameters by the user. The study involves the development of a graphical 2D mapping of Wi-Fi signal areas on the floors of the building, which will eventually allow to correctly geometrically visualize the coverage of points and determine the output of their signals beyond the perimeter of the building, which is potentially dangerous in terms of Wi-Fi network security.

Keywords: web application; client; server; Wi-Fi network; security; access point range; obstacles.

Введение

Веб-приложение по контролю формы периметра покрытия беспроводной сети разрабатывалось для визуализации зон покрытия Wi-Fi на схеме пользователя. Для этого пользователю необходимо загрузить в приложение схему помещений, установить на схеме роутеры, а также установить параметры роутеров и в конце указать все препятствия, которые должны быть учтены на схеме объекта [1]. Также веб-приложение может обрабатывать данные, полученные с помощью Acrylic Wi-Fi Heatmaps и отчетов Wi-Fi анализатора Fluke Aircheck и данные полученные со специальных устройств для измерения Wi-Fi сигнала.

1. Методология исследования

Методика расчета дальности действия Wi-Fi

Для того что бы узнать расчетное расстояние дальности действия точки доступа в первую очередь приведем инженерную формулу расчета потерь в свободном пространстве FSL (Free Space Loss), дБ:

$$FSL = 33 + 20\ (\lg F + \lg D), \quad (1)$$

где F– центральная частота канала, на котором работает система связи, МГц;

D – расстояние между двумя точками, км.

FSL определяется суммарным усилением системы Y, дБ. Оно вычисляется следующим образом:

$$Y = P_t + G_t + G_r - P_{min} - L_t - L_r, \quad (2)$$

где P_t, – мощность передатчика, дБ/мВт;

G_t – коэффициент усиления передающей антенны, дБ;

G_r – коэффициент усиления приемной антенны, дБ;

P_{min} – чувствительность приемника на данной скорости, дБ/мВт;

L_t – потери сигнала в коаксиальном кабеле и разъемах передающего тракта, дБ;

L_r – потери сигнала в коаксиальном кабеле и разъемах приемного тракта, дБ.

FSL вычисляется по формуле:

$$FSL = Y - SOM, \quad (3)$$

где SOM (System Operating Margin) – запас в энергетике радиосвязи, дБ;

Y – суммарное усиление системы, дБ.

SOM учитывает возможные факторы, отрицательно влияющие на дальность связи [2], такие как:

- температурный дрейф чувствительности приемника и выходной мощности передатчика;

- всевозможные атмосферные явления: туман, снег, дождь;

- рассогласование антенны, приемника, передатчика с антенно-фидерным трактом.

Подставив значения, получим итоговую формулу дальности связи:

$$D = 10^{\left(\frac{FSL}{20} - \frac{33}{20} - lgF\right)}. \quad (4)$$

Полученная формула используется для расчета максимальной дальности отрисовки Wi-Fi сигнала.

2. Реализация приложения

Для реализации веб-приложения использовался язык Golang и JavaScript. Язык Golang был выбран, так как с его помощью можно быстро реализовать серверную часть. JavaScript был выбран вместе с фреймворков React, так как реализовать клиентскую часть легче с помощью компонентного подхода.

Работа приложения начинается с процедуры входа:

Сторона клиента.

На стороне клиента данная функция была реализована в виде ссылке на end-point сервера. Пользователю нужно нажать на картинку в виде Google, как произойдет переадресация на Google Cloud.

Сторона сервера.

Для начала, чтобы было возможно использовать данный вид аутентификации нужно получить Credentials с Google Cloud. Также после получения credentials нужно указать URL. В рассмотренном примере пробной версии приложения он имеет значение “http://localhost:8080/auth/callback”.

Для работы с OAuth на сервере предусмотрены две функции:

- loginWithGoogle();

- callback().

Механизмы визуализации собранной статистики требуют последовательной работы с данными клиентской и серверной частей приложения.

Сторона клиента.

Функция RouterSettings обрабатывает данные, которые пользователь внес в характеристики роутеров, для этого пользователю необходимо активировать на устройство на схеме. Приложение выводит окно, куда необходимо ввести данные о параметрах устройства. После указания характеристик роутеров, пользователю необходимо указать масштаб изображе-

ния и препятствия, которые находятся на схеме. Для этого пользователю необходимо зайти в специальный режим рисования препятствий.

За визуализацию различных препятствий отвечает функция objectClickListener создания объектов, которые наносятся на схему.

После завершения настроек объектов, пользователю необходимо нажать на кнопку Submit. Задействуется функция handleUpload.

Структура функции handleUpload:

- изначально создается новый объект типа FormData. В этот объект будут записываться различные данные, которые необходимо отправить на сервер;

- генерируется текущая схема помещения предприятия со всеми препятствиями;

- импортируются характеристики роутеров и перемещаются в массив;

- отправляется запрос на сервер приложения.

После того как сервер обработает запрос он пришлет ответ, в который будет входить изображение с визуализированной сетью Wi-Fi.

Сторона сервера.

Для обработки запроса с клиента был разработан специальный метод handlerMap. Далее описана логика этого метода:

- изначально сервер прочитывает запрос и забирает оттуда токен. Токен хранится в переменной под названием authorization. Далее происходит узнавание пользователя путям преобразование токена в id и логин пользователя. За это все отвечает метод getUserId;

- сервер получает из запроса все характеристики роутеров. За это отвечает функция getValuesOfRouter. Данная функция сначала вытягивает из запроса всю информацию о роутерах из параметра data. Далее парсит информацию о роутерах в структуру под названием RequestRoutters. После чего происходит валидация данных и перенос новых данных в новую структуру RouterSettings;

- из запроса извлекается изображение для целевой сети. За это отвечает функция getImageFromContext;

- после получения информации о роутерах и их откликах, создается объект класса, которые будет заниматься визуализацией сети Wi-Fi на схеме пользователя. Для создания объекта передаётся путь файла, характеристики роутеров и путь к новому файлу с визуализированной сетью Wi-Fi. Методы, который будет визуализировать Wi-Fi сеть называется drawOnImage();

- после визуализации изображения сервер переводит изображение в массив байтов с помощью метода readFile;

- в результате формируется ответ клиенту, в который будет, входит готовое изображение.

Стоит обратить внимание на то, с какими данными будет работать метод по визуализации данных, данные хранятся в структуре:

- первое поле - это массив из структуры, в котором хранятся координаты роутеров;

- второе поле, это массив из настроек роутеров.

Чтобы визуализация Wi-Fi сети производилась корректно, нужно рассчитать дальность Wi-Fi сети. За расчет данных отвечает функция calculationOfValues. Для описания расчета предложен следующий алгоритм:

- сначала из структуры берутся данные о скорости и канале. Эти данные будут заменяться на чувствительность и центральную частоту соответственно;

- далее происходит расчет FSL и расчет дальности Wi-Fi сигнала, все формулы для реализации этого были описаны в начале статьи.

Также стоит учесть, что данный сервис рассчитывает данные на частоте 2.4 и 5 ГГц.

3. Результат работы приложения

В результате работы веб-приложения были получены отрисовки, которые представлены на рисунке. Замеры производились для изучения зон покрытия учебного корпуса №5 УО «ГГУ имени Ф.Скорины».

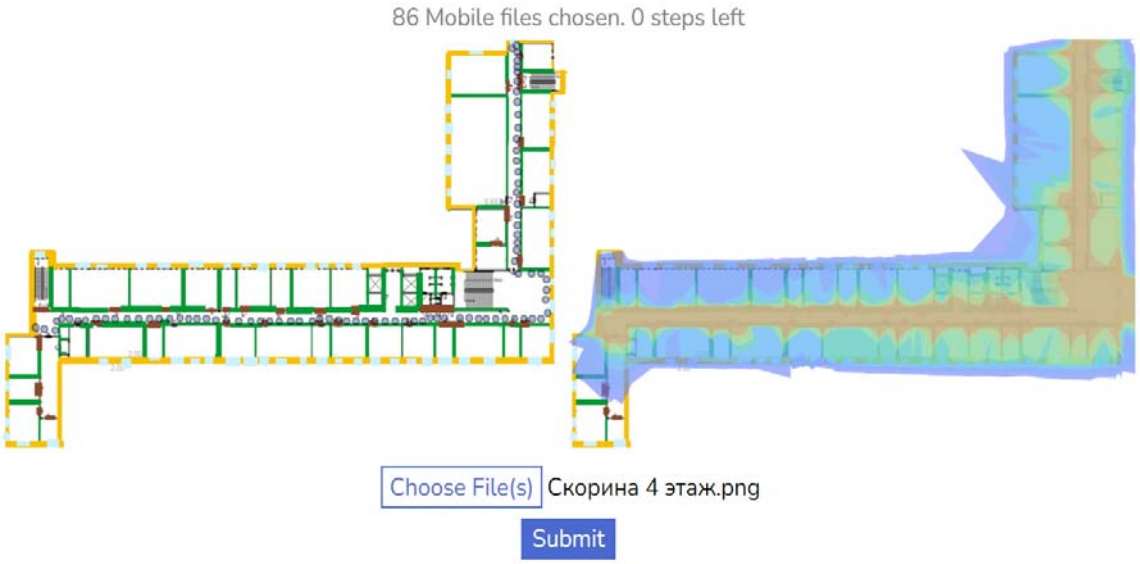


Рисунок – Выполнение отрисовки на базе данных Acrylic Wi-Fi Heatmaps

Заключение

Веб-приложение разработано для университета или других организаций, которым нужна безопасная Wi-Fi сеть, что на данный момент является весьма актуальной проблемой. Безопасной сетью, в частном случае, считается беспроводная сеть, зона вещания которой не выходит за пределы периметральных стен помещений предприятия. Разработанное приложение способно визуализировать зоны покрытия Wi-Fi сети на предварительно оцифрованном чертеже плане помещений и определять опасные зоны Wi-Fi сети.

Библиографические ссылки

1. Демиденко О.М., Кулинченко В.Н., Бычков П.В. Контроль и диагностика внутрипериметральных каналов независимых (смежных) беспроводных сегментов сети // Известия Гомельского государственного университета имени Ф. Скорины. 2021. № 6(129). С. 85-89.
2. Кулинченко В.Н., Демиденко О.М. Изучение влияния внешних помех на качество сигнала в сетях WI-FI // Проблемы физики, математики и техники. 2015. № 4(25). С. 96-99.

УДК 004.056

ИССЛЕДОВАНИЕ УЯЗВИМОСТЕЙ СИСТЕМНЫХ СТРУКТУР ЖЕСТКИХ ДИСКОВ

Т.Б. Ларина, М.Н. Падалка

ФБГОУ «Российский университет транспорта» (РУТ-МИИТ),
ул. Образцова, д.9, стр.9, 127994, г.Москва, Россия, tblarina@gmail.com, maksonmoskva@yandex.ru

В статье проводится анализ потенциальных уязвимостей системных структур жестких дисков компьютеров с MBR и GPT-конфигурацией. Показаны источники и последствия уязвимостей. Приводятся результаты экспериментов, демонстрирующие разрушительные последствия ряда уязвимостей. Анонсируются решения по организации защиты системных дисковых структур.

Ключевые слова: Компьютерная безопасность; защита данных; уязвимость дисков; системные структуры дисков; MBR; GPT.

INVESTIGATION OF VULNERABILITIES IN SYSTEM STRUCTURES OF HARD DRIVES

T.B. Larina, M.N. Padalka

Federal State Educational Institution "Russian University of Transport" (RUT-MIIT),9 Obraztsova str., 9, 127994, Moscow, Russia, tblarina@gmail.com, maksonmoskva@yandex.ru

The article analyzes the potential vulnerabilities of the system structures of hard drives of computers with MBR and GPT configuration. The sources and consequences of vulnerabilities are shown. The results of experiments are presented, demonstrating the devastating consequences of a number of vulnerabilities. Solutions are offered to organize the protection of system disk structures.

Keywords: Computer security; data protection; disk vulnerability; disk system structures; MBR; GPT.

Введение

Жесткий диск является основным устройством для размещения и функционирования операционной системы и долговременного хранения информации. Что делает его наиболее привлекательным компьютерным узлом для злоумышленников и основным источником проблем, связанных с доступом к данным. Авторы ставят своей целью провести исследование особенностей конфигураций жестких дисков и их системных структур с

точки зрения их уязвимости атакам типа «перехват управления» и «отказ в обслуживании».

1. Уязвимости системных структур жестких дисков традиционной конфигурации

Под традиционной конфигурацией имеют в виду MBR-конфигурацию жесткого диска. Для системных целей используется самый первый сектор жесткого диска - MBR-сектор. Он включает в себя код процедуры Главного загрузчика, Таблицу разделов и сигнатурный код [1]. Таблица разделов, состоящая из четырех записей, играет ключевую роль в процессе определения геометрии жесткого диска и загрузке с него операционной системы (ОС). Для преодоления ограничения на количество разделов предусмотрена возможность использовать «расширенный» тип раздела. Он позволяет создавать внутри себя любое количество внутренних разделов. Доступ к ним осуществляется по информации из системных EPR-секторов каждого внутреннего раздела.

Понимание процесса загрузки ОС с жесткого диска MBR-конфигурации важно для оценки проблем безопасности. Он состоит из трех этапов: «Диагностика», «Начальная загрузка» и собственно «Загрузка ОС» [1]. Этапы «Диагностика» и «Начальная загрузка» являются общими для всех ОС и выполняются процедурами системного сервиса базовой системы ввода-вывода (BIOS). На этапе «Диагностика» выполняется аппаратное самотестирование блока питания. Далее процедура BIOS Post выполняет начальное тестирование устройств. После завершения процедуры начинается второй этап загрузки. Третий этап «Загрузка ОС» является уникальным для каждой ОС.

В процессе исследований было выявлено две группы серьезных уязвимостей [2]. Это уязвимости, связанные с этапами загрузки ОС и самими системными структурами MBR-сектора:

- Обязательное чтение кода Главного загрузчика из MBR-сектора при включении компьютера и его общеизвестное размещение;
- Код Главного загрузчика исполняется в среде реального режима процессора, где отсутствуют аппаратные механизмы защиты;
- Поражение Таблицы разделов ведет к потере геометрии жесткого диска и соответственно, возможности работы с ним;
- MBR-дорожка жесткого диска не входит в разделы, невидима из файловых систем и может быть использована для размещения враждебного кода.

Таким образом, в результате уязвимости системных дисковых структур компьютер может быть подвержен двум типам атак:

- «Отказ в обслуживании» возникнет в результате уничтожения кода в MBR-секторе, что приведет к потере информации о конфигурации диска и невозможности загрузки ОС;
- «Перехват управления» может реализоваться размещением вредоносного кода в MBR-секторе и остальных секторах системной дорожки до запуска какой-либо антивирусной защиты.

2. Уязвимости системных структур жестких дисков с GPT конфигурацией

Ограничения классического BIOS стали создавать определенные проблемы его использования. Наиболее существенным является то, что 16-разрядные процедуры BIOS исполняются в реальном режиме процессора и ограничены адресным пространством памяти в 1 Мб. В результате совместной работы гигантов компьютерной индустрии появилась новая спецификация BIOS – Unified Extensible Firmware Interface (UEFI) [3]. Принципиальное ее отличие от классического BIOS состоит в разрядности исполняемого кода. Процедуры UEFI работают в 32-х разрядном защищенном режиме, что позволяет им использовать большие объемы оперативной памяти и аппаратно-программные механизмы защиты процессора. Появилась возможность организовать поддержку современных устройств и технологий.

Существенным изменениям подвергся и механизм загрузки операционных систем. В UEFI выделяют уже пять ключевых фаз загрузки: Security (SEC), Pre-EFI Initialization (PEI), Driver Execution Environment (DXE), Boot Device Selection (BDS) и Transient System Load (TSL) [4]. Уникальным для каждой операционной системы этапом является Transient System Load. Остальные стадии являются общими.

При включении компьютера начинается фаза Security, где выполняется первичная инициализация и перевод процессора в защищенный режим работы. Выполняется модуль инициализация Trusted Platform Module, который является корневой точкой доверия. Он осуществляет криптографическую проверку драйверов и EFI-приложений. В Pre-EFI Initialization выполняется инициализация процессора, его устройств окружения и оперативной памяти. Boot. Driver Execution Environment – загрузка всех необходимых драйверов. Boot Device Selection – аналог этапа «Начальная загрузка ОС» в традиционной конфигурации, его выполняет менеджер загрузки UEFI Boot manager. Этот модуль совмещает в себе функции традиционных BIOS – BootStrap и Главного загрузчика, и рас-

ширяет их функциональные возможности. Задача UEFI Boot manager - определить устройство загрузки ОС, выполнить загрузку EFI-драйверов для данных устройств, загрузить EFI приложения и стартовать уникальный пятый этап загрузки – Transient System Load [4].

В системных дисковых структурах произошли изменения (рис.1):

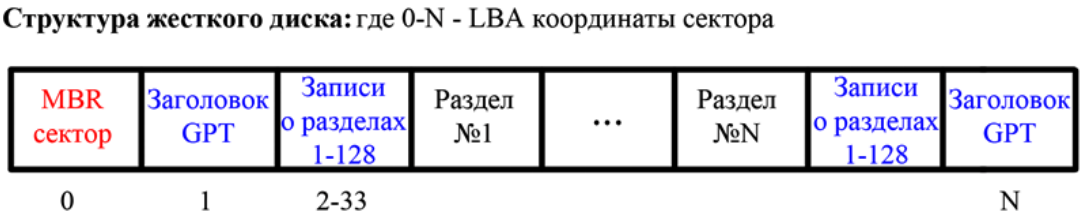


Рисунок 1 – Системные структуры жесткого диска с GPT

MBR-сектор больше не используется, он сохранен только для поддержки работы традиционных ОС. Создан новый тип таблицы разделов - GUID Partition Table и изменена структура записей о разделах. Она позволяет конфигурировать для работы жесткие диски емкостью до 8 Зетабайт. Новая структура Таблицы с увеличенным количеством записей позволило отказаться от использования «расширенного раздела». В конце жесткого диска расположена копия GPT-заголовка и самой таблицы, что позволяет обеспечить механизм проверки и восстановления структуры жесткого диска при различных системных сбоях [5].

Каковы потенциальные уязвимости в новых системных структурах?

1. Возможности проведения атаки «перехват управления»

Спецификация UEFI разрабатывалась с учетом высоких требований к безопасности. Существенно изменены механизмы загрузки операционных систем и системные дисковые структуры. При использовании ОС, ориентированных на работу с UEFI, опасность снижается благодаря следующим аспектам:

- ОС сразу начинает работать в 32-х или 64-х разрядном режиме, где существуют аппаратно-программные механизмы защиты;
- Механизм Trusted Platform Module накладывает дополнительные криптографические ограничения на исполняемый код;
- Начальную загрузку обеспечивает UEFI Boot manager, местонахождение которого на жестком диске неизвестно;
- Конфигурация диска определяется GPT-таблицей, которая поддерживает механизмы проверки на основе контрольных сумм, и существует копия данной таблицы в конце жесткого диска;
- Отсутствует структура расширенного раздела, создающая уже рассмотренные угрозы безопасности.

Все перечисленное существенно затрудняет какие-либо противоправные мероприятия по перехвату контроля над системой на общих этапах загрузки. Кроме того, для защиты от вредоносного кода на пятом этапе загрузки – TSL существует функция Boot Secure.

Тем не менее, система не стала полностью безопасной:

а) Новые механизмы способны защитить от полного перехвата управления только при работе в «идеальных условиях»: должна использоваться ОС, предназначенная для работы с UEFI и не должна отключаться функция Boot Secure. На практике эти условия часто не соблюдаются.

б) Для поддержки работы предшествующих версий ОС и ранних материнских плат в UEFI присутствует режим Legacy mode, создающий виртуальную среду «реального режима». В этом случае алгоритм загрузки ОС такой же, как в условиях классического BIOS. Следовательно, для режима Legacy mode характерны обе группы уязвимостей.

в) Несмотря на сложность внедрения вредоносного кода в этапы загрузки SEC, PEI, DXE и BDS, существует возможность программного внедрения на этапе TSL. Для защиты этого этапа предназначена опция Boot Secure, которая должна проверять EFI–приложения и загрузчики ОС с использованием криптографических алгоритмов. Проблема в том, что для проверки подлинности требуется подписать загрузчик и обеспечить наличие в постоянной памяти материнской платы проверочного сертификата. Это обстоятельство часто непреодолимо для производителей ОС и опцию Boot Secure отключают. Ситуация очень распространена в мире Unix-систем, где популярные дистрибутивы написаны энтузиастами и распространяются свободно.

2. Возможности проведения атак типа «отказ в обслуживании»

Встроенные в UEFI механизмы защиты GPT-таблицы базируются на том, что существует ее резервная копия и в MBR-секторе есть «защитная» запись. На первый взгляд этих изменений достаточно для обеспечения защиты от атак типа отказ в обслуживании. Однако наш эксперимент показывает, что это не так.

3. Эксперименты по исследованию последствий уязвимости

Проведены эксперименты, которые демонстрируют последствия разрушения системных структур дисков.

а) Уязвимость MBR-сектора в традиционных дисках

Разработана программа, осуществляющая атаку типа «отказ в обслуживании». Цель создания данной программы носит исключительно исследовательский характер. Для разработки программы использовался язык C++. Для получения необходимого уровня привилегий в программу

внедрен xml-манифест, запрашивающий выполнение кода в контексте администратора. Общий концепт программы таков. Создается нулевой массив размером в сектор, жесткий диск открывается, как текстовый файл, и нулевой массив записывается в MBR-сектор (рис.2).

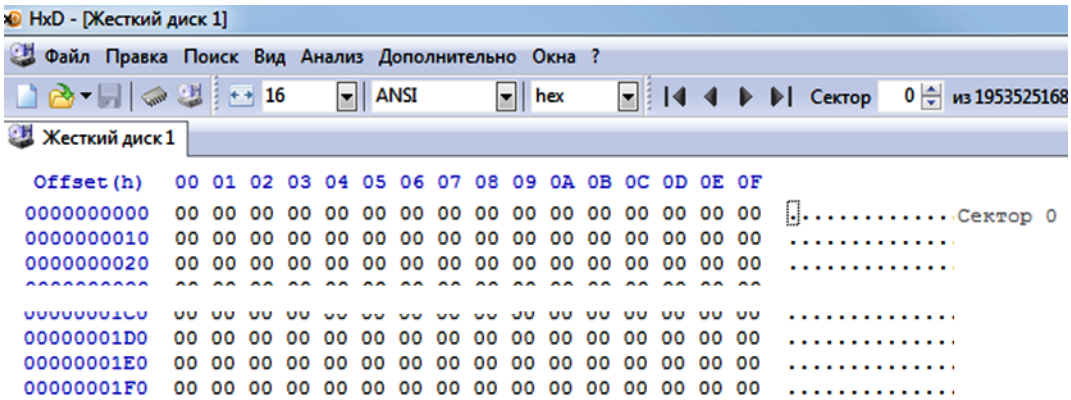


Рисунок 2 – Уничтоженный MBR-сектор

Диск программно закрывается и выполняется перезапуск компьютера командным интерпретатором. После чего полностью теряется доступ к ресурсам компьютера, поскольку ОС не может загрузиться с диска.

б) Уязвимость расширенного раздела.

При разрыве цепочки ссылок на внутренние разделы системных EPR-секторах расширенного раздела будет потерян доступ ко всем или отдельным внутренним разделам, в зависимости от места разрыва (рис.3).

Рисунок 3 – Уязвимость расширенного раздела

Для эксперимента была создана виртуальная машина с двумя жесткими дисками. Первый - системный раздел, на втором диске создан расширенный раздел с тремя внутренними разделами (рис.4).

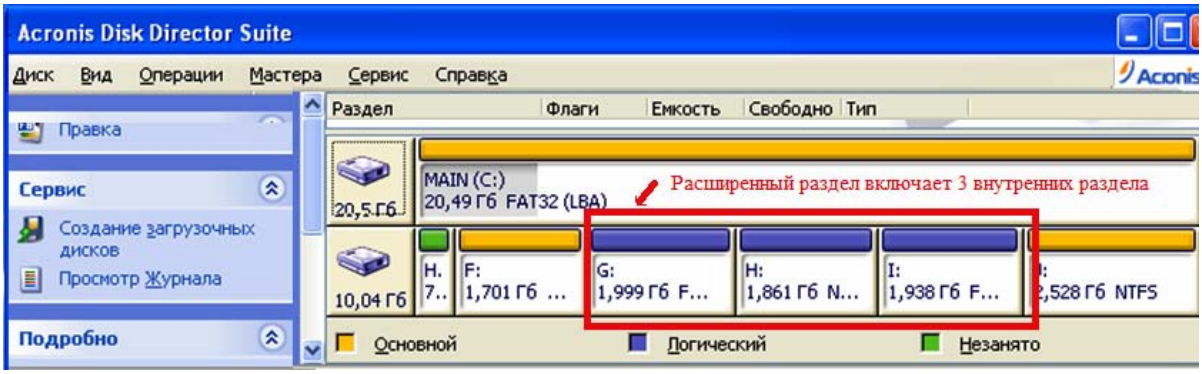


Рисунок 4 – Исходная конфигурация жестких дисков

Затем, в расширенном разделе был программно уничтожен первый EPR-сектор и произведена перезагрузка. В результате потери ссылок на логические диски, данные в расширенном разделе стали недоступны, раздел стал отображаться как неразмеченное пространство (рис.5).

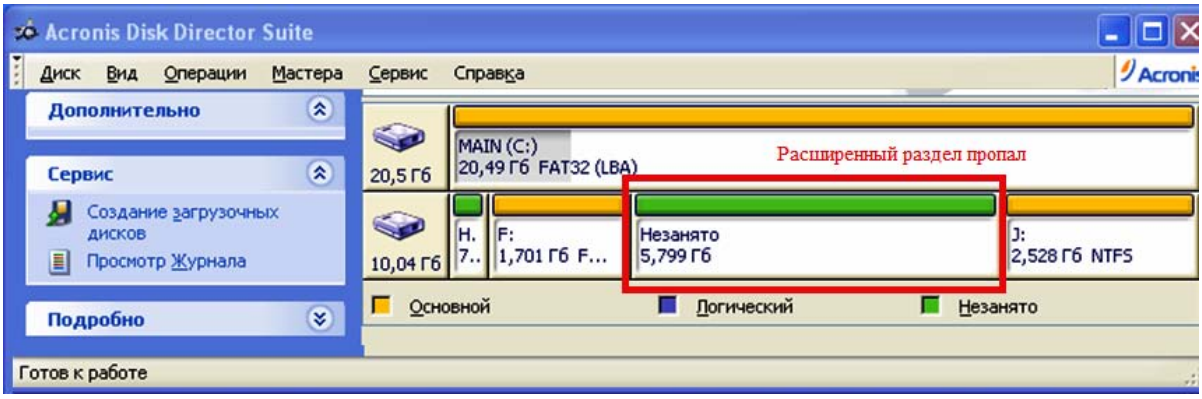


Рисунок 5 – Потеря расширенного раздела

Несмотря на то, что в базовых механизмах 32-х и 64-разрядных ОС есть запрет на низкоуровневый программный доступ к жесткому диску через прерывание INT13h BIOS, доступ к диску все же возможен через сервисы самой ОС. Во всех ОС существуют функции системного API, позволяющие открывать диск, как файл, и работать с его содержимым. Чем мы и пользовались в экспериментах.

в) Уязвимость GPT-таблицы разделов

Проведен эксперимент, показывающий, что атаку типа «отказ в обслуживании» на GPT-дисках осуществить не сложнее, чем на дисках с традиционной конфигурацией. Была разработана программа, которая используя системное API ОС, выполняет обнуление MBR-сектора, GPT-заголовка и существующих записей GPT-таблицы. После перезагрузки дисковый ресурс становится недоступным. При этом копия таблицы разделов в конце диска была проигнорирована самими средствами восста-

новления. Таким образом, несмотря на ряд серьезных нововведений в UEFI, данный тип атак по-прежнему осуществим.

Заключение

Проведенные эксперименты показывают, что системные дисковые структуры остаются уязвимыми для несанкционированного программного доступа. Гарантированно защитить их от потери информации можно только резервным копированием конфигурационной информации на внешние носители с возможностью восстановления. Это приводит к необходимости разработки программного средства, которое будет автоматически решать эту задачу вне зависимости от квалификации персонала. Разрабатывается специальный программный комплекс Disk Struct Saver, который позволяет выполнять автоматическое резервное копирование системных дисковых структур с MBR-конфигурацией в файл внешнего носителя и автоматического восстановления [6, 7].

Библиографические ссылки

1. Ларина Т.Б. Дисковые структуры операционных систем. Учебное пособие. М.: МИИТ, 2011. 173 с.
2. Ларина Т.Б., Падалка М.Н. Анализ уязвимостей системных дисковых структур операционных систем // Инновационные, информационные и коммуникационные технологии: сборник трудов XVII Международной научно-практической конференции / под. ред. С.У.Увайсова. Москва, Ассоциация выпускников и сотрудников ВВИА им. проф. Жуковского, 2020. С. 54–59.
3. Zimmer V, Marisetty S, Rothman M. Beyond BIOS. Developing with the Unified Extensible Firmware Interface. 3rd Edition. Walter de Gruyter Inc.,Boston, 2017. 356 с.
4. Интернет-ресурс: Блог по Windows / UEFI – Унифицированный расширяемый микропрограммный интерфейс. URL: http://datadump.ru/uefi/ (дата обращения: 29.06.2022).
5. Интернет-ресурс: FANDOM / OSDev Wiki / GUID-таблица разделов (GPT). URL: https://osdev.fandom.com/ru/wiki/GUID-таблица_разделов_(GPT) (дата обращения: 29.06.2022).
6. Ларина Т.Б., Падалка М.Н. Метод защиты и восстановления системных структур на жестких дисках // Инновационные, информационные и коммуникационные технологии: сборник трудов XVII Международной научно-практической конференции. /под. ред. С.У.Увайсова. Москва, Ассоциация выпускников и сотрудников ВВИА им. проф. Жуковского, 2020. С. 63–68.
7. Ларина Т.Б., Падалка М.Н. Разработка программного комплекса для защиты системных структур жестких дисков // Цифровые технологии и решения в сфере транспорта и образования: Материалы национальной научно-практической конференции. Мн.: Белый ветер, 2020. С. 87–94.

УДК 004.021:004.4

КОНТРОЛЬ ЦЕЛОСТНОСТИ И СООТВЕТСТВИЯ ВЕРСИЙ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ УПРАВЛЕНИЯ СИСТЕМАМИ ПЕРЕМЕЩЕНИЙ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

А.Ф. Марко

Белорусский государственный университет информатики и радиоэлектроники, ул. П. Бровки, 6, 220013, г. Минск, Беларусь, mmts@bsuir.by

Представлены методы и алгоритмы контроля целостности и соответствия версий программного обеспечения для управлении системами перемещений реального времени.

Ключевые слова: Программное обеспечение; соответствие версий; контроль целостности; системы перемещений; технология EtherCAT.

CONTROL OF INTEGRITY AND COMPLIANCE OF VERSIONS OF SOFTWARE FOR MANAGEMENT OF MOVEMENT SYSTEMS IN REAL TIME

A.F. Marko

Belarusian State University of Informatics and Radioelectronics, 6 P. Brovki Street, 220013, Minsk, Belarus, mmts@bsuir.by

Methods and algorithms for monitoring the integrity and compliance of software versions for managing real-time motion systems are presented.

Keywords: Software; Version Compliance; Integrity Control; Motion Systems; EtherCAT technology.

Введение

Объединение узлов точной механики с электронными, электрическими и компьютерными компонентами позволило осуществлять проектирование и производство качественно новых модулей, систем и машин с их интеллектуальным управлением. С развитием электрических приводов и возможностей их применения в индустриально-производственных и

транспортных системах, стала очевидна необходимость полной интеграции составляющих элементов электропривода: механики, электрических машин, силовой электроники, микропроцессорной техники и программного обеспечения для наиболее полного использования возможностей современного электропривода, и построения на его основе мехатронных систем перемещения [1].

Проведённый анализ современных программно-аппаратных средств показал, что наиболее эффективной технологией для реализации управления системами многокоординатных перемещений в режиме реального времени является технология EtherCAT, внедрение которой требует разработки дополнительного программного обеспечения. Разработка такого программного обеспечения выполняется с применением специальных инструментов, которые повышают эффективность разработки за счёт снижения трудоёмкости выполняемых операций. К таким инструментам относятся различные среды разработки программного обеспечения и системы контроля версий. В данных инструментах существует проблема отсутствия универсального решения для версионирования dll-библиотек и исполняемых exe-файлов программного обеспечения с целью установления связи между данными файлами и их исходным кодом. Также недостаточно проработан вопрос контроля целостности программного обеспечения, особенно объектов базы данных [1].

В рамках настоящей работы рассматривается программное обеспечение для контроля целостности и соответствия версий при управлении системами многокоординатных перемещений в режиме реального времени.

1. Теоретические основы

Для управления в реальном времени всё большее распространение получает технология EtherCAT. Из всех устройств, подключённых к шине EtherCAT, только мастер может быть инициатором телеграмм. Все остальные устройства модифицируют проходящую через них телеграмму, читая и записывая в неё данные технологического процесса. Аппаратная задержка на прохождение телеграммы через одно slave-устройство составляет всего несколько наносекунд [1].

Рассматриваемая в работе EtherCAT-сеть (рисунок 1) содержит один управляющий компьютер master и шесть локальных систем управления, каждая из которых работает в режиме slave и обеспечивает реализацию прецизионных перемещений соответствующего планарного позиционера по двум координатам.

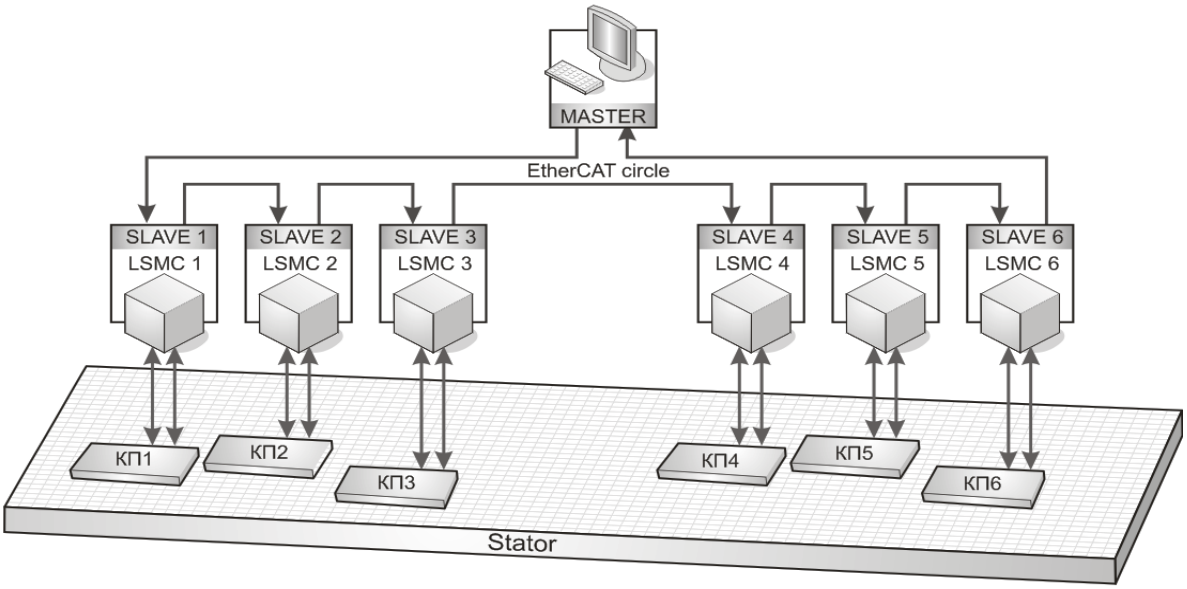


Рисунок 1 – Схема EtherCAT-сети для управления шестью позиционерами

Таким образом, EtherCAT-технология предоставляет разработчикам систем управления технологическими процессами и сложным оборудованием полностью интегрированное решение, обеспечивающее стандартную и надёжную сеть обмена управляющей информацией. При этом количество задействованных полевых шин и интерфейсов уменьшается, обеспечивая тем самым унификацию всех процессов управления, гибкость структуры при практически неограниченном количестве устройств и малое время реакции на события, а также обеспечивается возможность переконфигурирования системы управления без необходимости её полного отключения [1].

В связи с наметившимся внедрением технологии EtherCAT в прецизионное технологическое оборудование актуальной и важной является разработка специальных инструментов, позволяющих разрабатывать программное обеспечение системы управления в множестве версий и тем самым с постоянным изменением кода. Поэтому актуальной и важной является задача автоматизации контроля за соответствием версий компонентов такого программного обеспечения в процессе его разработки и контроля целостности в процессе эксплуатации.

2. Результаты и их обсуждение

Контроль за соответствием версий позволяет решить задачу обновления версий сборок с расширениями dll и exe при изменении их исходного кода, который компилируется в данные сборки при помощи среды Visual Studio (VS). Алгоритмы обновления версий реализованы в виде

расширения для среды VS, которая в свою очередь может взаимодействовать как с централизованной системой управления версиями Team Foundation Server (TFS), так и с децентрализованной системой Git [2]. Пользовательский интерфейс расширения встроен непосредственно в интерфейс среды VS, что позволяет контролировать соответствие версий и разрабатывать программное обеспечение в одном окружении. Данное расширение определяет какие компоненты программного обеспечения изменены по отношению к последней версии в системе TFS или Git, формирует новую версию, присваивает данную версию компонентам и сохраняет изменения в систему TFS или Git.

В случае программного обеспечения, разрабатываемого на языке C#, базовыми компонентами являются так называемые проекты [2]. В процессе разработки была реализована концепция формирования версий для проектов, как принадлежащих к версионируемому решению, так и для проектов, подключённых из других решений по ссылке. В свою очередь в проектах, принадлежащих к версионируемому решению, выделяются основные проекты, которые являются источником версии последнего релиза. Определение типа проекта выполняется с помощью структурного анализа файла решения с расширением sln и его конфигурационного файла. Версия проекта состоит из двух частей: ручной части (первые три старших разряда версии), определяемой последней версией релиза и автоматической части, соответствующей номеру сохранения в системе TFS или Git, в котором был изменён проект. Каждый проект содержит текстовый файл AssemblyInfo, который хранит версию проекта. Для её получения или изменения используются регулярные выражения. В результате после сборки релиза формируются dll- и exe-файлы с актуальными версиями [3, 4].

Алгоритмы формирования и сравнения контрольных сумм в процессе эксплуатации встроены непосредственно в программное обеспечение системы управления и предназначены для определения незапланированных изменений [3, 4]. Программное обеспечение для системы управления состоит из множества различных объектов, таких как исполняемые файлы, файлы данных и объекты баз данных, формирование контрольных сумм выполняется для каждого типа по-разному. Также принимается во внимание, что некоторые объекты, такие как таблица пользователей или регистрационные файлы изменяются в процессе эксплуатации, следовательно, контрольные суммы для них не формируются. На рисунке 2 приведены выделяемые типы и многоступенчатость процесса формирования контрольных сумм.

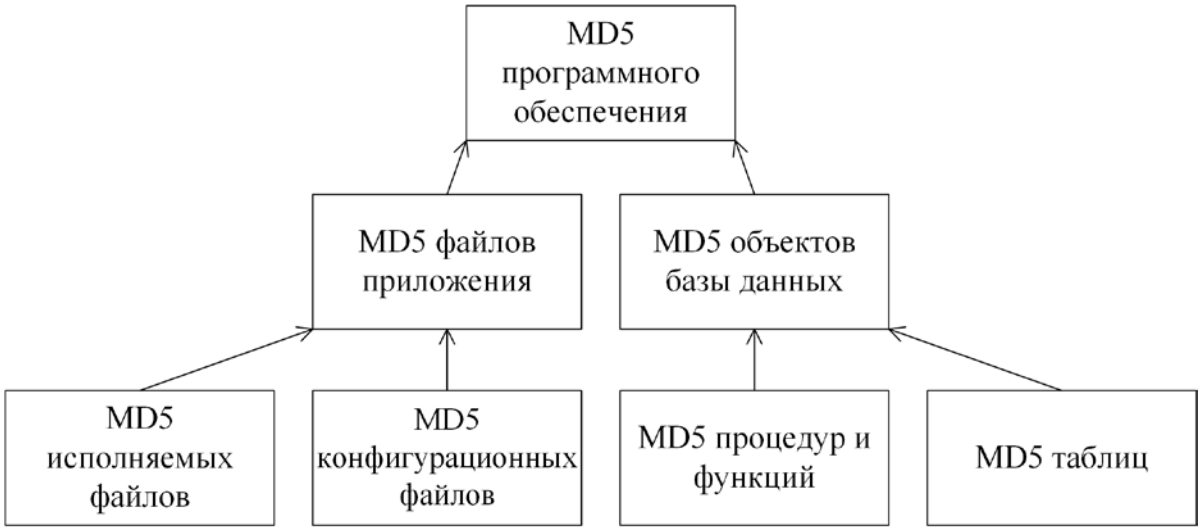


Рисунок 2 – Многоступенчатый процесс формирования контрольных сумм

Основная цель многоступенчатого формирования контрольных сумм заключается в удобстве представления информации о состоянии каждой подгруппы объектов в отдельности и всей системы в целом, а также в сокращении времени, необходимого для нахождения изменённых объектов.

Таким образом были разработаны: метод и алгоритмы контроля за соответствием версий компонентов программного обеспечения, заключающиеся в автоматизированном обновлении версий dll-библиотек и исполняемых exe-файлов при внесении изменений в их исходный код, а также метод и алгоритмы контроля целостности программного обеспечения, заключающиеся в формировании эталонных контрольных сумм с использованием хэш-функции MD5 для объектов программирования и объектов баз данных, сравнении их с текущими контрольными суммами, и позволяющие детектировать любые изменения указанных объектов и тем самым уменьшить вероятность использования программного обеспечения с незапланированными изменениями. Данные методы и алгоритмы оказалось весьма востребованными для систем управления реального времени на многокоординатных приводах прямого действия, таких как тестеры печатных плат, сборочное и оптико-механическое оборудование микроэлектроники.

Библиографические ссылки

1. Карпович С.Е. Системы многокоординатных перемещений на механизмах параллельной кинематики. Минск: Бестпринт, 2017. 254 с.

2. Шарп Дж. Microsoft Visual C#. Подробное руководство. 8-е изд. СПб.: Питер, 2017. 848 с.
3. Марко А.Ф., Чеушев К.В., Лобашинский М.В. Программное средство для обеспечения целостности при разработке и эксплуатации системы автоматизированного управления транспортным оборудованием // Информационные технологии и системы 2018 (ИТС 2018): материалы международной научной конференции. 25 октября 2018. Минск: 2018. С. 122–123.
4. Марко А.Ф., Кузнецов В.В., Войтов А.Ю. Программный модуль контроля целостности в системах управления реального времени // BIG DATA и анализ высокого уровня: сборник материалов V Международной научно-практической конференции. 13–14 марта 2019. Минск: 2019. С. 221–223.

УДК 519.719.2

РАЗДЕЛЕНИЕ СЕКРЕТА В ПОСТКВАНТОВЫЙ ПЕРИОД

Г.В. Матвеев

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, matveev@bsu.by

Криптография на основе решеток в настоящее время является одной из самых популярных областей математической криптографии. Криптографические конструкции на основе решеток являются ведущими кандидатами для постквантовой криптографии с открытым ключом. Это еще больше мотивирует изучение криптографических конструкций на основе решеток. В этой статье мы предлагаем метод разделения секрета основанный на теории решеток.

В разделе 1 содержатся необходимые сведения из постквантовой криптографии, краткий обзор двух работ по разделению секрета, а также постановка задачи и формулировка цели и задач исследования.

В разделе 2 приведены основные факты по теории модулярного разделения секрета и указаны ссылки на работы по теории решеток, на которых основано наше исследование.

В разделе 3 содержаться полученные результаты, их обсуждение и сравнение с уже известными результатами.

Ключевые слова: постквантовая криптография; криптография на основе решеток; китайская теорема об остатках; модулярное разделение секрета.

SECRET SHARING IN THE POST-QUANTUM PERIOD

G.V. Matveev

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus,
e-mail: matveev@bsu.by

Lattice-based cryptography is one of the most popular areas in mathematical cryptography nowadays. Lattice-based cryptographic constructions are the leading candidates for public-key post-quantum cryptography. This further motivates the study of lattice-based cryptographic constructions . In this paper, we introduce a method of lattice-based secret sharing scheme.

The organization of the paper is as follows:

Section 1 contains the necessary information from post-quantum cryptography, a brief overview of two works on the secret sharing, as well as the formulation of the task and the formulation of the purpose and objectives of the study. Section 2 provides basic facts on the theory of modular secret sharing and references on the theory of lattices on which our study is based. Section 3 contains the results obtained, their discussion and comparison with the known results.

Keywords: post-quantum cryptography; lattice-based cryptography; Chinese Remainder Theore; modular secret sharing.

Введение

Идея квантовых вычислений была независимо предложена Юрием Ивановичем Маниным и Ричардом Фейнманом в начале 1980-х. С тех пор была проделана большая работа по созданию действующего квантового компьютера.

Известно, что квантовый компьютер может значительно ускорить решение ряда задач, таких как факторизация чисел и дискретное логарифмирование в группе точек эллиптической кривой. Это становиться существенной проблемой для криптографии, так как безопасность распространенных стандартизированных систем зависит от сложности решения этих задач.

Тем не менее, квантовые вычисления довольно длительное время оставались лишь потенциальной возможностью, которую нельзя было технически реализовать. Однако в последнее время перспектива создания квантовых компьютеров улучшилась [1] и это стимулировало NIST объявить открытый конкурс по созданию новых постквантовых стандартов. Основное требование для алгоритмов шифрования постквантовой криптографии состоит в том, что они должны быть основаны на NP- сложных задачах. Отметим такие знаковые события события:

2003 год: Д.Бернштейн предлагает термин “постквантовая криптография”,

2006 год: первая конференция PQCrypto,

2017 год: объявлен конкурс NIST.

Уже проведены три этапа конкурса NIST, определены перспективные направления и произведен конкурсный отбор участников.

В настоящее время разработка алгоритмов постквантовой криптографии ведется по четырем направлениям, использующих:

1. Теорию решеток (Lattice based cryptography),
2. Коды, исправляющие ошибки (Code based cryptography),
3. Многочлены в конечных полях (Multivariate, quadratic equations cryptography),
4. Теорию хэш-функций для больших данных (Hash-based cryptography).

На третьем этапе конкурса NIST победителями признаны:

1. По направлению, использующему теорию решеток: три криптосистемы с открытым ключом CRYSTALS-Kyber, NTRU, SABER и две системы цифровой подписи CRYSTALS-Dilithium, FALCON.

2. По направлению, использующему коды, исправляющие ошибки – криптосистема Мак-Элиса.

3. По направлению, использующему многочлены в конечных полях – система цифровой подписи Rainbow.

По итогам конкурса напрашивается очевидный вывод о том, что теория решеток становится вычислительной базой основных постквантовых стандартов. Это, безусловно, повлечет разработку новых и оптимизацию известных решеточных алгоритмов, что делает привлекательным и естественным более широкое криптографическое применение этих алгоритмов.

В настоящей работе теория решеток применяется и для изучения модулярного разделения секрета. Первые результаты в этой области получил Н. Шенец в работе [2], в которой рассматривается задача построения модулярных схем разделения секрета на основе целочисленных решеток.

Особенностью подхода, предложенного Н. Шенцом, является использование мономиального упорядочения на полугруппе $Z^n > 0$, которое применяется для определения частичного секрета участника. В рамках предложенного подхода была построена однородная асимптотически идеальная многомерная пороговая схема разделения секрета.

Решеточный подход применяется и для решения еще одной важной задачи в теории разделения секрета. А именно, в работе [3] предложен способ увеличения порога для любой схемы Шамира даже в случае если такое увеличение и не предполагалось заранее. Найденное решение поставленной задачи не требует дополнительного обмена информацией между дилером и участниками протокола. Надо сказать, что эта проблема решалась и ранее, но лишь путем дополнительного обмена информацией между дилером и участниками протокола либо для этой цели разрабатывались специальные схемы разделения секрета. Основе предложенного способа лежит решеточное декодирование кода Рида-Соломона.

В настоящей работе предлагается более простой способ разделения секрета не использующий мономиальное упорядочение для определения частичного секрета участника. Получен ряд структурных результатов и, в частности, построена однородная асимптотически идеальная многомерная пороговая схема разделения секрета.

1. Методология исследования

Для решения поставленной задачи используется классическая теория сравнений (модулярная арифметика) и CRT-алгоритм, в частности, а также теория решеток [4]. Полученные результаты можно рассматривать как многомерное обобщение модулярного целочисленного разделения секрета [5].

Напомним основные понятия и задачи теории разделения секрета.

Под *схемой разделения секрета* (СРС) понимают распределение секрета c на части $c_1, c_2, ..., c_t$ между t участниками и такой алгоритм вычисления секрета c, при котором его могут вычислить лишь заранее определенные (разрешенные) подмножества участников.

Дадим теперь строгие определение структуры доступа и пороговой структуры доступа в частности.

Определение 1. *Под структурой доступа Г будем понимать любое семейство подмножеств множества всех участников со свойством монотонности, т.е.*

$$A \in \Gamma, A \subset B \subset I \Rightarrow B \in \Gamma.$$

Среди СРС важное место занимают пороговые схемы. СРС называется *(k,t)*-пороговой схемой, если разрешенными являются все подмножества мощности не меньше k, где k - некоторое фиксированное число $1 \le k \le t$.

Под реализацией структуры доступа Г будем понимать такой алгоритм, который позволяет восстанавливать секрет лишь для подмножеств участников, содержащихся в семействе Γ.

При разработке схем разделения секрета стараются удовлетворить нескольким естественным требованиям. К их числу в первую очередь относится требование *идеальности* схемы разделения секрета, т.е. чтобы размер частичного секрета c_i не превышал размера основного секрета c. С другой стороны, желательно, чтобы неразрешенные множества участников не получали никакой дополнительной информации к имеющейся априорной о возможном значении секрета c – это требование *совершенности*. Иногда требование *идеальности* включает в себя требование *совершенности.*

В криптографии традиционно широко используется вычисление в кольцах вычетов Z_m. Видимо, этим объясняется популярность следующей пороговой *(k,t)*-схемы. Рассмотрим систему $m_1 < m_2 < ... < m_t$ попарно взаимно простых модулей, для которых выполнено условие

$$M_1 = m_1 m_2 ... m_k > m_{t-k+2} m_{t-k+3} ... m_t = M_2.$$

Одновременно требуется, чтобы разность $M_1 - M_2$ была по возможности большой. Секрет c выбирается случайным образом из промежутка (M_2, M_1), а частичный секрет c_i i-го участника, $i = 1, 2, ..., t$, есть наименьший неотрицательный вычет c по модулю m_i. Предполагается, что каждый участник знает не только свой частичный секрет c_i, но и модуль m_i.

В основе модулярной схемы лежит утверждение о том, что любая система сравнений

$$\begin{cases} x \equiv c_{i_1} (\mathrm{mod}\, m_{i_1}), \\ x \equiv c_{i_2} (\mathrm{mod}\, m_{i_2}), \\ \quad \vdots \\ x \equiv c_{i_s} (\mathrm{mod}\, m_{i_s}) \end{cases}$$

имеет единственное решение в промежутке (M_2, M_1), если $s \geq k$, и имеет достаточно много решений в противном случае.

Основная трудность в построении этих схем заключается в подборе модулей $m_1, m_2, ..., m_t$, удовлетворяющих условию $M_1 = m_1 m_2 ... m_k > m_{t-k+2} m_{t-k+3} ... m_t = M_2$. По этой причине в кольце целых чисел условия идеальности и совершенности можно реализовать лишь с некоторым приближением. Как показано в работах [6], [7] этот недостаток схемы устраняется путем перехода от кольца целых чисел к кольцу многочленов от одной переменной над полем Галуа. На этом пути впервые были построены идеальные модулярные схемы разделения секрета.

2. Результаты и их обсуждение

Рассмотрим конечнопорожденный Z-модуль Z^n. Пусть $a_1, a_2, \ldots, a_n$ линейно независимые над полем R векторы из Z^n.

Определение 2. *Решеткой Λ в Z^n называется множество всех векторов (точек) $x = \sum_{i=1}^{n} u_i a_i$, где $u_i \in Z$, а векторы $a_1, a_2, \ldots, a_n$ называют базисом решетки Λ.*

Если каждая точка решетки Λ является также точкой решетки M, то Λ называется подрешеткой решетки M.

Для построения схемы разделения секрета необходимо определить:

1. как выбирать модули участников (здесь в качестве модулей выступают подрешетки),
2. как строить вектор-вычет (частичный секрет) участника,
3. как выбирать вектор-секрет для заданной структуры доступа,
4. как восстанавливать вектор-секрет по частичным секретам.

В работе [2] предложен способ построения частичного секрета использующий мономиальные упорядочения. Сначала на $Z^n > 0$ задается

порядок. Каждому участнику i дается в качестве открытого ключа некоторая подрешетка Λ_i, заданная ее базисной матрицей A^i. Секретом является некоторая точка $c \in Z^n > 0$.

Как и в одномерном случае, каждый участник в качестве частичного должен получить минимальный в некотором смысле представитель своего класса $\{A^i x + c,\ x \in Z^n\}$. А именно, среди всех представителей класса $\{A^i x + c,\ x \in Z^n\}$ участнику i дается минимальный относительно заданного порядка на $Z^n > 0$ вектор c_i. Так определяется процедура приведения секрета по модулю подрешетки [2].

Для восстановления секрета с необходимо найти пересечение классов, а именно решить следующую задачу:

$$\{A^i x + s^i\} \cap \{A^j y + s^j\} = \{A^{ij} z + s^{ij}\},\ x, y, z \in Z^n,$$

где A^{ij} – базисная матрица пересечения подрешеток Λ_i и Λ_j, а s^{ij} минимальный представитель пересечения классов. Ясно, что в общем случае пересечение классов может быть пусто. В данном случае пересечение не пусто, поскольку каждому классу принадлежит точка с.

Отметим, что выбор (генерация) модулей и секрета в предложенной схеме напрямую зависит от рассматриваемой структуры доступа и заданного порядка.

Наш подход основан на следующем важнейшем параметре решетки и не зависит от мономиальных упорядочений решеток.

Определение 3. Фундаментальным параллелепипедом решетки называется множество точек

$$P = \{\sum_{i=1}^{n} x_i\ a_i, 0 \le x_i\ < 1\}.$$

1. В качестве открытых ключей участников выбираются подрешетки $\{\Lambda_i\}$, $i=1,2,\ldots,n$.

2. Частичный секрет участника определяется как вычет вспомогательного секрета $S = \sum_{j=1}^{j=n} \alpha_j a_j$, $\alpha_j \in R$ по модулю подрешетки $s_i = S \bmod \Lambda_i$. Под вычетом в данном случае понимается представитель смежного класса $S + \Lambda_i$ в фундаментальном параллелепипеде решетки Λ_i

$$s_i = \sum_{j=1}^{j=n} \{\alpha_j^i\} a_j,\ 0 \le \{\alpha_j^i\} < 1,$$

где $\{\alpha_j^i\}$ – дробная часть α_j^i.

В рамках предложенного подхода получены следующие результаты.

Теорема 1. *Существует решеточно-модулярная реализация произвольной структуры доступа.*

Теорема 2. *Существует однородная асимптотически совершенная и асимптотически идеальная пороговая схема разделения секрета в* Z^n.

Качество схемы можно улучшить путем перехода от целочисленной решетки Z^n к решетке над кольцом многочленов от одной переменной над полем Галуа.

Теорема 3. *Существует идеальная и совершенная решеточно-модулярная реализация пороговой структуры доступа в решетке над кольцом многочленов.*

Замечание. *Предложенный алгоритм для вычисления вычета по модулю решетки имеет полиномиальную сложность* $O(n^3)$.

Таким образом, к настоящему времени в статьях [2], [3] и в данной работе решены следующие задачи:

1. Предложены алгоритмы приведения секрета по модулю подрешетки.

2. Доказана возможность многомерной модулярной реализации произ-вольной структуры доступа.

3. Построена однородная асимптотически совершенная и асимптотически идеальная пороговая схема разделения секрета в Z^n.

4. Решена задача по увеличению порога схемы разделения секрета после распределения частичных секретов

Библиографические ссылки

1. Bernstein D., Lange T. Post-quantum cryptography // Nature, № 549. 2017. P. 188–194. URL: https://doi.org/10.1038/nature23461.
2. Шенец Н.Н. Многомерное модулярное разделение информации // Информатика. 2007. № 4(16). С. 125–132.
3. Steinfeld R., Pieprzyk J. and Wang H.. Lattice-Based Threshold Changeability for Standard Shamir Secret-Sharing Schemes // IEEE Transactions on Information Theory, vol. 53, no. 7, p. 2542-2559, July 2007, doi: 10.1109/TIT.2007.899541.
4. Касселс Дж.В.С. Введение в геометрию чисел. М.: Мир, 1965. 421 с.
5. Asmuth C.A., Bloom J. A modular approach to key safeguarding // IEEE Trans. on inf. theory. 1983. Vol.29. P. 156–169.
6. Galibus T, Matveev G. Generalized Mignotte's Sequences Over Polynomial Rings Electronic Notes // Theoretical Computer Science. 2007. Vol. 186. P. 43–48. DOI: 10.1016/j.entcs.2006.12.044.
7. Galibus T, Matveev G, Shenets N. Some structural and security properties of the modular secret sharing. In: Symbolic and Numeric Algorithms for Scientific Computing, 2008. SYNASC 2008. // Los Alamitos: IEEE Computer Society Press. 2009. P. 197–200. DOI: 10.1109/SYNASC.2008.14.

УДК 004.4: 519.23

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ЭНТРОПИЙНОГО АНАЛИЗА ДИСКРЕТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

В.Ю. Палуха, Ю.С. Харин, М.В. Мальцев, А.И. Сергеев, А.А. Орлов

НИИ прикладных проблем математики и информатики БГУ
Пр-т. Независимости, 4, 220030, Минск, Беларусь
palukha@bsu.by, kharin@bsu.by, giftis95@mail.ru,
alex.orlov.official@gmail.com, maltsew@bsu.by

Рассматривается задача анализа дискретных последовательностей на основе оценок функционалов энтропии Шенннона, Реньи и Тсаллиса. Представлен разработанный программный комплекс.

***Ключевые слова*:** функционалы информационной энтропии; энтропия Шеннона; энтропия Реньи; энтропия Тсаллиса; статистические оценки.

SOFTWARE COMPLEX FOR ENTROPY ANALYSIS OF DISCRETE SEQUENCES

U.Yu. Palukha, Yu.S. Kharin, M.U. Maltsau, A.I. Siarheeu, A.A. Arlou

RI for Applied Problems of Mathematics and Informatics, BSU
4 Niezalieznasci Avenue, Minsk 220030, Belarus
Corresponding author: palukha@bsu.by

The problem of analyzing discrete sequences based on the estimates of the Shannnon, Renyi, and Tsallis entropy functionals is considered. The developed software package is presented.

Keywords: information entropy; Shannon entropy; Renyi entropy; Tsallis entropy; statistical estimators.

Введение

Стойкость систем криптографической защиты информации зависит от того, насколько близка используемая ими случайная или псевдослучайная последовательность по своим свойствам к равномерно распределённой случайной последовательности (РРСП) [1] (которая также называется «чисто случайной»), что устанавливается с помощью статистических тестов. В них проверяется гипотеза $H_* = \{\{x_t\}$ является РРСП$\}\}$. В данной работе в качестве тестовых статистик выступают статистические оценки

энтропии Шеннона, Реньи и Тсаллиса. Авторами разработан программный комплекс, который позволяет вычислять оценки указанных функционалов энтропии дискретной последовательности и на их основе принимать или отклонять гипотезу о «чистой случайности» анализируемой последовательности.

Энтропийный анализ

Пусть на вероятностном пространстве (Ω, F, P) с множеством состояний $\Omega = \{\omega_1, \ldots, \omega_N\}$ определена случайная величина $x = x(\omega) = \omega$ с дискретным распределением вероятностей $p = \{p_k\}$, $p_k = P\{x = \omega_k\}$, $p_k \geq 0$, $\sum_{k=1}^{N} p_k = 1, k = 1, \ldots, N$. В таблице приведены формулы наиболее распространённых функционалов энтропии.

Таблица – Функционалы энтропии

Энтропия Шеннона	$H(p) = -\sum_{i=1}^{N} p_i \ln p_i$
Энтропия Реньи	$H_r(p) = \frac{1}{1-r} \ln\left(\sum_{i=1}^{N} p_i^r \right), r \in \square, r > 1.$
Энтропия Тсаллиса	$S_r(p) = \frac{1}{r-1}\left(1 - \sum_{i=1}^{N} p_i^r \right), r \in \square, r > 1.$

Пусть наблюдается реализация случайной последовательности $\{x_t : t = 1, \ldots, n\}$ длины n из распределения вероятностей $\{p_k\}$, по которой будет оцениваться энтропия. Частотные оценки вероятностей имеют вид

$$\hat{p}_k = \frac{v_k}{n}, \quad v_k = \sum_{t=1}^{n} I\{x_t = \omega_k\} \in \square_0 = \square \cup \{0\}, \quad I\{x_t = \omega_k\} = \begin{cases} 1, x_t = \omega_k; \\ 0, x_t \neq \omega_k. \end{cases} \quad (1)$$

Рассмотрим асимптотику

$$n, N \to \infty, n/N \to \lambda, 0 < \lambda < \infty. \quad (2)$$

которая отличается от классической ($n \to \infty$, $N < \infty$) тем, что длина последовательности n и мощность алфавита N растут синхронно.

Оценка энтропии Шеннона на основе статистик (1) имеет вид:

$$\hat{H} = \hat{H}(n, N) = -\sum_{k=1}^{N} \hat{p}_k \ln \hat{p}_k = -\sum_{k=1}^{N} \frac{v_k}{n} \ln \frac{v_k}{n} = \ln n - \frac{1}{n} \sum_{k=1}^{N} v_k \ln v_k. \quad (3)$$

Теорема 1 [2]**.** В асимптотике (2) статистика (3) при гипотезе H_* имеет асимптотически нормальное распределение с параметрами

$$\mu_H = \ln n - e^{-\lambda}\sum_{k=1}^{+\infty}\frac{\ln(k+1)\lambda^k}{k!}, \tag{4}$$

$$\sigma_H^2 = \frac{e^{-\lambda}}{n}\sum_{k=1}^{+\infty}\frac{(k+1)\lambda^k}{k!}\ln^2(k+1) - \frac{e^{-2\lambda}}{N}\left(\sum_{k=1}^{+\infty}\frac{\ln(k+1)\lambda^k}{k!}\right)^2 - \\ -\frac{e^{-2\lambda}}{n}\left(\sum_{k=1}^{+\infty}\ln(k+1)\frac{\lambda^k}{k!}(k+1-\lambda)\right)^2. \tag{5}$$

Из теоремы 1 видно, что в асимптотике (2) оценка (3) является смещённой. Для функционалов энтропии Реньи и Тсаллиса можно построить несмещённую оценку в асимптотике (2), в т.ч. и при $\lambda < 1$.

Определим r-ую нисходящую факториальную степень x:

$$x^{\underline{r}} = x(x-1)\ldots(x-r+1) = \frac{x!}{(x-r)!} = \sum_{i=0}^{r} s(r,i)x^i, \tag{6}$$

где $s(r, i)$ – число Стирлинга первого рода; при $x < r$ полагают $x^{\underline{r}} ::= 0$.

Статистические оценки энтропии Реньи и Тсаллиса, построенные с использованием нисходящей факториальной степени, имеют вид

$$\hat{H}_r(n, N) = \frac{1}{1-r}\ln\left(\sum_{k=1}^{N}\frac{\nu_k^{\underline{r}}}{n^r}\right) = \ln n + \frac{1}{r-1}\left(\ln n - \ln\sum_{k=1}^{N}\nu_k^{\underline{r}}\right), \tag{7}$$

$$\hat{S}_r(n, N) = \frac{1}{r-1}\left(1 - \sum_{k=1}^{N}\frac{\nu_k^{\underline{r}}}{n^r}\right) = \frac{1}{r-1}\left(1 - \frac{1}{n^r}\sum_{k=1}^{N}\nu_k^r\right). \tag{8}$$

Теорема 2 [2]**.** В асимптотике (2) статистика (8) является состоятельной асимптотически несмещённой оценкой энтропии Тсаллиса и при истинной гипотезе H_* имеет асимптотически нормальное распределение с параметрами:

$$\mu_{S,r} = \frac{1}{r-1}\left(1 - \frac{1}{N^{r-1}}\right), \tag{9}$$

$$\sigma_{S,r}^2 = \frac{\lambda^{r-1}}{(r-1)^2 n^{2r-1}}\left(\sum_{i=1}^{r} s(r,i)\sum_{j=1}^{i-1} C_i^j r^{i-j}\sum_{k=1}^{j} S(j,k)\lambda^k - r^2\lambda^{r-1} + r!\right), \tag{10}$$

где $S(r, i)$ – число Стирлинга второго рода.

Следствие 1. При $r = 2$ для математического ожидания и дисперсии асимптотического распределения оценки (8) справедливы выражения:

$$\mu_{S,2} = 1 - \frac{1}{N}, \quad \sigma^2_{S,2} = \frac{2}{Nn^2}.$$

Теорема 3 [2]**.** В асимптотике (2) статистика (7) является состоятельной оценкой энтропии Реньи и при истинной гипотезе H_* имеет асимптотически нормальное распределение с параметрами:

$$\mu_{H,r} = \ln N, \tag{11}$$

$$\sigma^2_{H,r} = \frac{\sum_{i=2}^{r} s(r,i) \sum_{j=1}^{i-1} C_i^j r^{i-j} \sum_{k=1}^{j} S(j,k) \lambda^k - r^2 \lambda^{r-1} + r!}{(r-1)^2 n \lambda^{r-1}}. \tag{12}$$

Следствие 2. При $r = 2$ для дисперсии асимптотического распределения вероятностей оценки (7) справедливо выражение:

$$\sigma^2_{H,2} = \frac{2}{n\lambda}.$$

Пусть $\alpha \in (0,1)$ – заданный уровень значимости. Введём обозначения: $\hat{h}$ – статистическая оценка энтропии Шеннона (3), Реньи (7) или Тсаллиса (8), μ_h – асимптотическое математическое ожидание статистической оценки энтропии Шеннона (4), Реньи (11) или Тсаллиса (9), σ_h^2 – асимптотическая дисперсия статистической оценки энтропии Шеннона (5), Реньи (12) или Тсаллиса (10) при истинной гипотезе H_*. Решающее правило имеет вид [2]:

$$\text{принимается} \begin{cases} H_*, \text{если} \quad t_- < \hat{h} <; t_+ \\ \overline{H_*}, \text{в противном случае,} \end{cases} \qquad t_\pm = \mu_h \pm \sigma_h \Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \tag{13}$$

где $\Phi(\cdot)$ – функция распределения стандартного нормального закона.

Вычислим нормированную статистику

$$\tilde{h} = \frac{\hat{h} - \mu_h}{\sigma_h}.$$

Она в асимптотике (2) имеет стандартное нормальное распределение: $\tilde{h} \sim \mathcal{N}(0,1)$. Следовательно, двустороннее *p*-значение для неё равно

$$p - value = 2\left(1 - \Phi\left(\left|\tilde{h}\right|\right)\right). \tag{14}$$

Пусть генератор порождает двоичную выходную последовательность $\{y_\tau\}$, $\tau = 1, \ldots, T$. «Нарежем» её на непересекающиеся подряд идущие фрагменты длины s (s-граммы): $X^{(t)} = (X_j^{(t)}) = (y_{(t-1)s+1}, \ldots, y_{ts}) \in \{0,1\}^s$, $t = 1, \ldots, n = [T / s]$. Из полученных s-грамм сформируем новую последовательность $\{x_t\}$ из алфавита мощности $N = 2^s$ по правилу $x_t = \sum_{j=1}^{s} 2^{j-1} X_j^{(t)} + 1$.

На основе критерия (13) мы можем вычислить последовательность нормированных отклонений оценки энтропии от математического ожидания в зависимости от s, которые назовём **энтропийными профилями**:

$$\chi(s) = \frac{\hat{h}(s) - \mu_h(s)}{\sigma_h(s)\Phi^{-1}(1-\alpha/2)} = \frac{\tilde{h}(s)}{\Phi^{-1}(1-\alpha/2)}, \; s = s_-, \ldots, s_+. \qquad (15)$$

Программный комплекс

В НИИ ППМИ разработан программный комплекс (ПК), который позволяет вычислять оценки энтропии Шеннона (3), Реньи (7) и Тсаллиса (8) при $r = 2$, их асимптотические параметры распределений вероятностей при гипотезе H_* с помощью алгоритмов [3], p-значения (14) и энтропийные профили (15) для двоичных файлов. Помимо вывода самих значений, программа выводит графики зависимостей этих величин от длины фрагмента s.

В начале работы необходимо выбрать файл с последовательностью, диапазон $s_-, \ldots, s_+$ и функционалы энтропии. Вычисляемые значения добавляются на экран в режиме реального времени. Имеется возможность изменять уровень значимости $\alpha \in (0,1)$ без пересчёта оценок энтропии и переключаться на различные режимы отображения: непосредственно оценки энтропии $\hat{h}$, нормированные значения (15), p-значения (14).

Для тестирования ПК подготовлена библиотека последовательностей псевдослучайных и физических генераторов. На рисунке 1 представлен результат работы ПК с последовательностью физического генератора [4], на рисунке 2 – с последовательностью, полученной при помощи регистра сдвига с линейной обратной связью (РСЛОС) с примитивным характеристическим многочленом над полем GF(2) $x^{32} + x^{22} + x^2 + x + 1$ [1] на уровне значимости $\alpha = 0.05$. Как видно из рисунков, для физического генератора гипотеза H_* принимается, для РСЛОС начиная с $s = 16$ отклоняется.

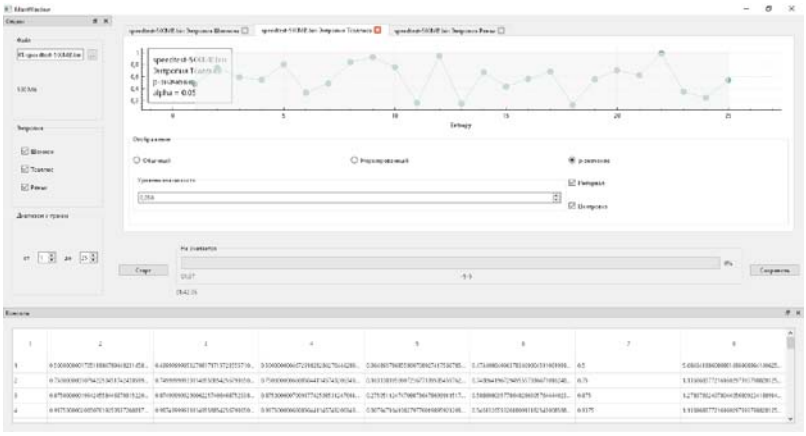


Рисунок 1 – Энтропийный профиль Тсаллиса физического генератора

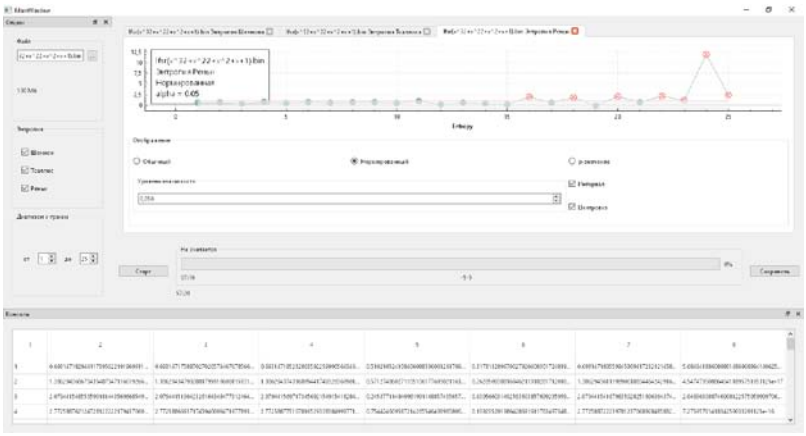


Рисунок 2 – Энтропийный профиль Реньи РСЛОС

Библиографические ссылки

1. Харин Ю.С., Агиевич С.В., Васильев Д.В., Матвеев Г.В. Криптология. Минск: БГУ, 2013. 512 с.
2. Палуха В. Ю. Статистические тесты на основе оценок энтропии для проверки гипотез о равномерном распределении случайной последовательности // Весці НАН Беларусі. Серыя фізіка-матэматычных навук. 2017. № 1. С.: 79–88.
3. Палуха В.Ю., Харин Ю.С. Вычисление статистических оценок функционалов энтропии двоичных последовательностей // Международный конгресс по информатике: информационные системы и технологии [Электронный ресурс]: Материалы международного научного конгресса. Республика Беларусь, Минск, 24–27 октября 2016 года. Минск: БГУ, 2016. С. 472–476.
4. Физический генератор. URL: http://qrng.physik.hu-berlin.de/files/speedtest-500MB.bin.

УДК 004.056:621.37

ПРОБЛЕМЫ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ЭЛЕКТРОМАГНИТНЫХ ИЗЛУЧЕНИЙ И НАВОДОК В СРЕДСТВАХ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

И.А. Третьяков, Я.И. Рушечников

ГОУ ВПО «Донецкий национальный университет», ул. Университетская, 24, 283001, г. Донецк, Донецкая Народная Республика,
i.tretiakov@mail.ru, ya.rushechnikov@donnu.ru

В данной работе рассмотрена проблема информационной безопасности средств вычислительной техники в виде побочных электромагнитных излучений и наводок. Рассмотрены процедуры обнаружения технических каналов утечки информации, визуализации радиообстановки и определения частот, на которых может быть перехвачена информация.

***Ключевые слова*:** ПЭМИН; программно-определяемая радиосистема; радиообстановка; обнаружение информации; быстрое сканирование.

PROBLEMS OF INFORMATION SECURITY OF ELECTROMAGNETIC RADIATION AND LEADS IN COMPUTER EQUIPMENT

I.A. Tretiakov, IA. I. Rushechnikov

Donetsk National University, 24 Universitetskaya str., 283001, Donetsk, Donetsk People's Republic, i.tretiakov@mail.ru, ya.rushechnikov@donnu.ru

In this paper, the problem of information security of computer equipment in the form of side electromagnetic radiation and interference is considered. The procedures for detecting technical channels of information leakage, visualizing radio placement and determining the frequencies at which information can be intercepted are considered.

Keywords: TEMPEST; software-defined radio system; radio placement; information detection; fast scanning.

Введение

В настоящее время широко известно, что работа структурных элементов электронных вычислительных устройств сопровождается побочными электромагнитными излучениями [1-3]. Это приводит к появлению наводок (в цепях проводных линий передачи, питания, заземления и т.д.) вследствие электромагнитного воздействия. Электромагнитные излучения, источником которых являются элементы и устройства вычислительной техники, как правило, должны отвечать нормам электромагнитной

совместимости. Однако, они не являются безопасными с точки зрения сохранения конфиденциальности обрабатываемой информации, откуда следует, что требования к электромагнитной совместимости и защите информации не являются взаимоисключающими.

Проблема безопасности побочных электромагнитных излучений и наводок (ПЭМИН) известна со времен появления самих средств электронной вычислительной техники. Она заключается в том, что информацию, обрабатываемую средствами вычислительной техники, можно восстановить путем приема и обработки побочных электромагнитных излучений и наводок. Применение в средствах вычислительной техники импульсных сигналов прямоугольной формы и высокочастотной коммутации приводит к тому, что в спектре из лучений будут компоненты с частотами вплоть до сверхвысокочастотного диапазона. Хоть и энергетический спектр таких сигналов убывает с повышением частоты, при этом увеличивается эффективность излучения и уровень излучений может оставаться постоянным до частот порядка нескольких гигагерц. Также резонансы из-за паразитных связей могут вызывать усиление излучения на некоторых частотах спектра.

Целью настоящей работы является исследование возможностей обнаружения побочных электромагнитных излучений и наводок в средствах вычислительной техники как источников информации.

Средства обнаружения ПЭМИН

Одним из эффективных средств обнаружения технических каналов утечки информации и побочных электромагнитных излучений и наводок в средствах вычислительной техники являются программно-определяемые радиосистемы (SDR) [4, 5]. В рамках поставленной задачи авторами применен комплекс программно-определяемых радиосистем HackRF One.

Особенность выбранного устройства заключается в том, что его возможности развиваются не только за счёт того, что пишется специальное программное обеспечение, но и благодаря обновлению микропрограммы самого устройства. Благодаря такой возможности устройство приобретает новые функции, которые изначально не были в него заложены. Таким образом развивается и возможность быстрого сканирования, которая позволяет охватить частотный диапазон от 1 до 6000 МГц менее чем за одну секунду. Данная способность порождает несколько проблем:

– по умолчанию приёмник получает снимок радиочастотной обстановки на весь свой диапазон, но с шагом перестройки в 1

МГц, что составляет 5999 выборок. Это не позволяет обнаруживать частоты, находящиеся внутри промежутка 1 МГц.

– при увеличении разрешающей способности сканирования (шага перестройки) приёмника значительно увеличивается объём обрабатываемых данных растёт (в два раза на каждом шаге). Это приводит к потребности в дополнительных аппаратных ресурсах вычислительного устройства, к которому подключается программно-определяемая радиосистема.

В таблице представлены данные зависимости количества элементов результирующей выборки от ширины окна просмотра программно-определяемой радиосистемы при быстром сканировании.

Таблица – Зависимость количества отсчётов от ширины окна при сканировании диапазона от 1 до 6000 МГц

Ширина окна, (Гц)	Количество отсчётов
1047552	6000
523776	13200
261888	22800
130944	46800
65472	92400
32736	183600
16368	368400

Из анализа таблицы обнаружена обратная зависимость, при которой чем уже ширина окна для просмотра, тем больше данных приёмник сгенерирует в результирующей выборке, а это в свою очередь может повлиять на скорость обработки и визуализации полученной информации.

Визуализация радиообстановки. Для получения наиболее точной картины радиообстановки и поиска частот вещания ПЭМИН была предложена следующая методика измерения:

– производится снимок радиообстановки вне зоны действия ПЭМИН вычислительной техники,

– приёмник с антенным трактом переносится ближе к вычислительной технике и эксперимент повторяется.

– дополнительно визуализируется значение среднего по измерениям (две параллельные линии), и чем оно больше, тем больше разница в радиообстановке при измерениях.

Как видно из рисунка 1, на некоторых частотах есть явное превышение мощности у одного спектра. Для получения наиболее явной разност-

ной картины проведено ещё одно измерение на интересующих участках. Результат этого измерения представлен на рисунке 2.

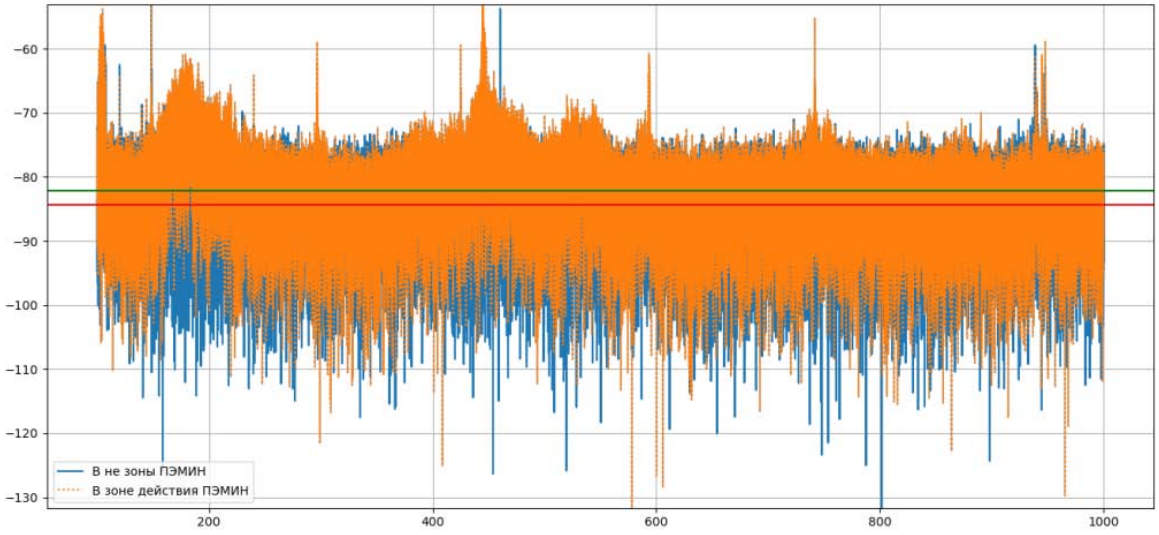


Рисунок 1 – Визуализация радиообстановки
на частотах от 1 до 1000 МГц

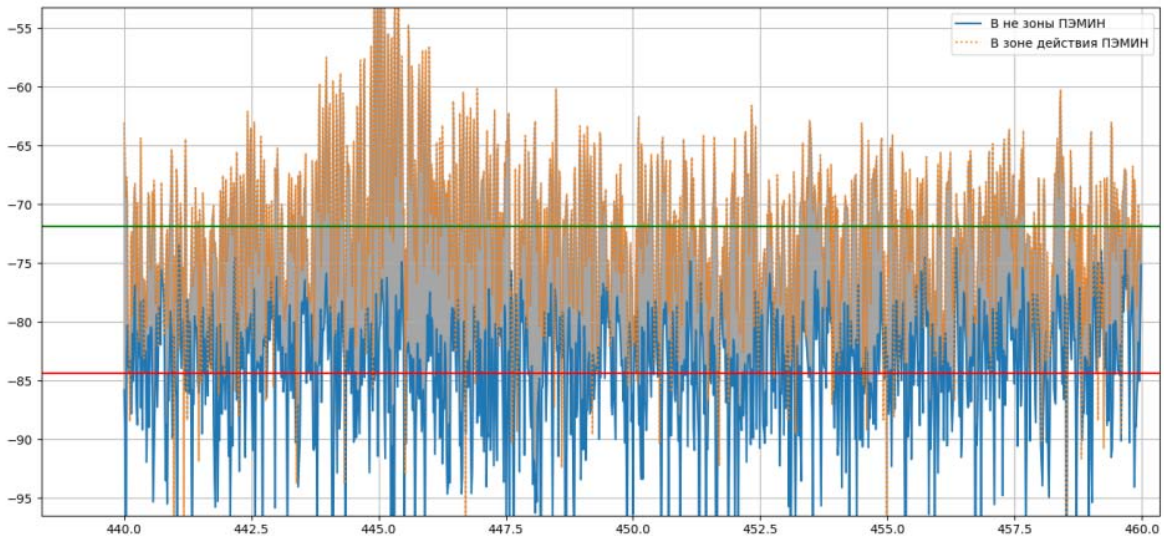


Рисунок 2 – Визуализация радиообстановки
на частотах от 440 до 460 МГц

По полученным разностным картинам (рис. 1, рис. 2) можно судить о наличии частот, на которых во время эксперимента было обнаружено несоответствие относительной мощности вещания. Следовательно, к данным частотным участкам можно применять дальнейшие процедуры для анализа и распознавания информации.

Заключение

Таким образом, рассмотрена проблема информационной безопасности побочных электромагнитных излучений и наводок в средствах вычислительной техники в рамках обнаружения технических каналов утечки информации, визуализации радиообстановки и определения частот, на которых может быть перехвачена информация.

Также рассмотрена возможность быстрого сканирования радиообстановки и установлена зависимость объема данных в результирующей выборке от ширины окна просмотра программно-определяемой радиосистемы.

Библиографические ссылки

1. Хорев А.А. Оценка возможности обнаружения побочных электромагнитных излучений видеосистемы компьютера // Доклады ТУСУР. 2014. №2(32). С. 207–213.
2. Антипов Д.А. Анализ утечек информации на основе побочных электромагнитных излучений // Доклады ТУСУР. 2018. №2. С. 27–32. DOI: 10.21293/1818-0442-2018-21-2-27-32.
3. Рушечников Я.И., Яновский А.В., Жинкина А.С., Данилов В.В. Электромагнитные излучения элементов электронной вычислительной техники // Вестник Донецкого национального университета. Серия Г: Технические науки. 2019. № 2. С. 25–35.
4. Рушечников Я.И., Данилов В.В. Информационная технология радиомониторинга на основе программно-определяемой радиосистемы // Вестник Донецкого национального университета. Серия Г: Технические науки. 2020. № 1. С. 31–36.
5. Третьяков И.А., Данилов В.В. Исследование спектрограмм радиочастот методами лингвистического анализа // Вестник Астраханского государственного технического университета. Серия: управление, вычислительная техника и информатика. 2020. № 3. С. 26-33. DOI: 10.24143/2072-9502-2020-3-45-51.

УДК [004.056+003.26](075.8)

КОРРЕКЦИЯ ОДИНОЧНЫХ И ДВОЙНЫХ ПАРНЫХ ОШИБОК В СТЕГАНОГРАФИЧЕСКИХ КАНАЛАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

П.П. Урбанович[1, 2]

[1]*Белорусский государственный технологический университет, ул. Свердлова, 13а, 220005 Минск, Беларусь,* p.urbanovich@belstu.by
[2]*Люблинский Католический университет им. Яна Павла II, al. Racławickie 14, 20-950 Lublin, Poland,* pavel.urbanovich@kul.pl

Проанализированы особенности использования избыточных корректирующих кодов в стеганографических приложениях. Приведено формальное описание модели стеганосистемы, в которой код используется для коррекции извлекаемого из стеганоконтейнера сообщения. Предложена конструкция линейного кода для коррекции одиночных и двойных парных (смежных) ошибок.

Ключевые слова. Линейный корректирующий код; парная ошибка; стеганография; скрытый канал.

CORRECTION OF SINGLE AND DOUBLE ERRORS IN STEGANOGRAPHIC CHANNELS OF INFORMATION TRANSMISSION

P.P. Urbanovich[1, 2]

[1]*Belarusian State Technological University, Sverdlova str., 220005 Minsk, Belarus,* p.urbanovich@belstu.by
[2]*The John Paul II Catholic University of Lublin, Al. Racławickie 14, 20-950 Lublin, Poland,* pavel.urbanovich@kul.pl

The features of the use of redundant correcting codes in steganographic applications are analyzed. A formal description of the steganosystem model, in which the code is used to correct the message extracted from the steganocontainer, is given. A construction of a linear code for correction of single and double paired (adjacent) errors is proposed.

Keywords. Linear error-correcting code; pair error; steganography; covert channel.

Введение

Стеганографические алгоритмы позволяют скрывать конфиденциальную информацию в каналах, формируемых носителями другой информации, или стеганографическими контейнерами (c). Такие каналы называют скрытыми. Исходная скрываемая информация (m), помимо основных стеганографических трансформаций (осаждение/извлечение), может

подвергаться также другим канальным преобразованиям: например, на основе криптографии и/или помехоустойчивого кодирования. Перечисленные преобразования информации на обеих сторонах канала можно соотнести с многоключевой стеганосистемой, основу которой составляет скрытый канал или стеганосообщение (*s*): сообщение + контейнер [1].

Идея использования помехоустойчивых кодов совместно со стеганографией в наиболее общем виде изложена в [2]; здесь автор назвал метод кодирования матричным. Одна из практических реализаций матричного кодирования связана с известным стеганографическим алгоритмом F5 [3]. Дальнейшие исследования в указанной предметной области преследовали по существу одну основную цель: обеспечить минимальные искажения *c* при осаждении исходного сообщения *m*, ориентируясь, как правило, на алгоритмы из класса LSB (List Sygnificant Bits – наименее значащих битов). Для решения этой задачи были созданы коды для записи на «мокрой бумаге», называемые также кодами «мокрой бумаги» (Wet Paper Codes, WPC) [4], комбинации кодов Хемминга и WPC [5], синдромные решетчатые коды (Syndrome Trellis Codes, STC) [6], стеганографические полярные коды (Steganographic Polar Codes, SPC) [7] и др.

Адаптация битов сообщения *m* из множества *M* под содержание соответствующих битовых последовательностей контейнера *c* из множества *C* на основе перечисленных или подобных им корректирующих кодов повышает стеганостойкость системы передачи или хранения скрытого сообщения. В данном случае под стеганостойкостью понимается степень модификации исходного содержания *c* после размещения в нем *m*, т. е. после преобразования контейнера *c* в стеганоконтейнер *s*: $c \rightarrow s$ ($s \in S$, *S* – множество всех возможных типов контейнеров).

Однако не менее важной является проблема обеспечения целостности осажденного в контейнер сообщения *m* после случайной или преднамеренной модификации стеганоконтейнера *s*. Нами, в частности, установлено, что простейшие конвертации *s* (docx–pdf–docx) в случаях использования *c* в виде текстового документа при реализации метода LSB на основе модификации битов цветовых каналов модели RGB при осаждении битов сообщения *m* в некоторых случаях приводит в конечном итоге к изменению цветового кода (в границах от 0 до 255). Если при осаждении *m* модифицируются два младших бита (из восьми) во всех или в отдельных цветовых кодах (R, G, B) пикселей, составляющих *c*, то с наибольшей вероятностью в результате указанных конвертаций значения одного или обоих битов, составляющих часть *m*, могут измениться. Это означает, что формальное воздействие на скрытый канал помехи (в виде указанной конвертации или в ином проявлении) приводит к появлению одиночной, а также двойной смежной (или парной)

ошибки. Практически единственным противодействием ошибкам является корректирующий код.

Специфика ошибок диктует необходимость поиска такой кодовой конструкции, которая обеспечивала бы положительный эффект в сравнении с применением известных кодов, обнаруживающих и корректирующих одиночные и двойные независимые ошибки в кодовых словах подобно тому, как может решаться задача повышения функциональной надежности систем и устройств полупроводниковой памяти [8]. Рассмотрению и анализу такого кода и посвящена настоящая работа.

1. Основная часть.

1.1. Формальное описание стеганографической системы и ее элементов.

Рассматриваемая стеганографическая система S с использованием корректирующего кода в наиболее общем виде может быть представлена следующим образом [1]:

$$S = (M, C, S, K, K_д, F, F^{-1}), \quad (1)$$

где $M, C, S, K, K_д$ – соответственно конечные множества, содержащие: возможные тайные сообщения M ($m \in M$); используемые контейнеры ($c \in C$); стеганоконтейнеры ($s \in S$); основные ключи (K, $k \in K$; k – отдельно взятый ключ), относящиеся к используемым стеганографическим методам (например, LSB) осаждения/извлечения тайного сообщения; дополнительные ключи ($K_д$, $k_д \in K_д$; $k_д$ – отдельно взятый дополнительный ключ), относящиеся к модификациям основного стеганометода (например, количество младших битов кодов R, G, B, используемых для внедрения m; или порядок выбора элементов c – последовательно или по иному принципу – для размещения битов m). При этом полагаем, что конкретное сообщение m в двоичном виде состоит из t битов: $m = m_1, m_2, \ldots, m_t$, которые разделяются на b блоков по l битов в каждом: $t = bl$.

Последние элементы в правой части (1): F, F^{-1}– функциональные преобразования (отображения), соответственно обозначающие внедрение закодированных на основе избыточного кода сообщений в контейнер и обратные функциональные преобразования: извлечение закодированных сообщений и исправление в них обнаруженных ошибок. В общем случае

$$F: M \times f_{ECC} \times C \times K \times K_д \rightarrow S, \quad (2)$$

$$F^{-1}: S' \times K \times K_д \rightarrow M' \times f_{ECC}, C. \quad (3)$$

В последних выражениях произведение $M \times f_{ECC}$ означает выполнение операции избыточного кодирования каждого из блоков сообщений M на

основе кода, корректирующего ошибки (Error Correcting Code, ECC); произведение $S' \times K \times K_д$ – функцию извлечения сообщений M' из стеганоконтейнеров S', $M' \times f_{ECC}$ – декодирование извлеченных сообщений M' с обнаружением и исправлением ошибок. Полагаем, что различные вышеупомянутые воздействия на скрытый канал стеганосистемы **S** (стеганоконтейнер S, $s \in S$) проводят к модификации S: $S \rightarrow S'$, $S \neq S'$, соответственно $s \neq s'$ ($s' \in S'$) и $M \neq M'$; в конечном итоге $m \neq m'$, где $m' = (m')_1, (m')_2, \ldots, (m')_t$.

1.2. Избыточный помехоустойчивый код, корректирующий одиночные и двойные парные ошибки.

Дальнейший анализ будем производить, опираясь на общепринятые положения теории избыточного кодирования информации (см., например, [9]).

Любой корректирующий код задается тремя основными параметрами: длиной кодируемого или информационного слова (в нашем случае – l), общей длиной кодового слова или блока (обозначим его символом n) и минимальным расстоянием Хемминга (d) между двумя кодовыми словами: (n, l, d)-код. Отдельный блок исходного сообщения m обозначим $\dot{m}_L$ (m состоит из b таких блоков); $\dot{m}_L = \dot{m}_1, \dot{m}_2, \ldots, \dot{m}_l$ и $\dot{m}_i = \{1, 2\}$, $i = 1, \ldots, l$. Закодированное сообщение $\dot{m}_L$ состоит из n символов. Это закодированное сообщение обозначим $\dot{m}_N$ ($\dot{m}_N = \dot{m}_1, \dot{m}_2, \ldots, \dot{m}_l, \dot{m}_{l+1}, \dot{m}_{l+2}, \dot{m}_{l+r}$; $n=l+r$).

Избыточность кода – $R = n/l$. Применительно к одному блоку внедряемого в контейнер s сообщения кодовое слово $\dot{m}_L$ является результатом выполнения операции $M \times f_{ECC}$ в выражении (2).

Подобным образом будем обозначать части извлекаемого из стеганоконтейнера s' закодированного блока (обозначим его $\dot{m}_N'$; $\dot{m}_N' = \dot{m}_1', \dot{m}_2', \ldots, \dot{m}_l', \dot{m}_{l+1}', \dot{m}_{l+2}', \dot{m}_{l+r}'$) сообщения.

Основная часть нашего исследования и анализа состоит в разработке такой конструкции линейного блочного (длина блока или кодируемого сообщения $\dot{m}_L$ равна l битам) корректирующего кода, который позволял бы обнаруживать и корректировать одиночные и двойные ошибки в смежных символах кодового слова $\dot{m}_N'$. Эта операция использования кода формально описывается частью выражения (3): $M' \times f_{ECC}$.

Линейный код Хемминга однозначно задается с помощью проверочной матрицы Хемминга, $\boldsymbol{H}_{r \times n}$ размерностью $r \times n$:

$$\boldsymbol{H}_{r \times n} = [\, \mathbf{I} \mid \mathbf{A} \,], \qquad (4)$$

где **I** – диагональная матрица размерностью $r \times r$, **A** – матрица размерностью $r \times l$, вес Хемминга каждого вектор-столбца $\mathbf{a}_j$ ($j = 1, 2, \ldots, l$) не менее двух: $wt(\mathbf{a}_j) \geq 2$.

При фиксированных l корректирующие свойства кода определяются параметром $R = (l+r)/l$, т.е. количеством r избыточных символов, которые нужно вычислить и присоединить к информационному слову $\dot{m}_L$.

Наши рассуждения при конструировании кода с заданными корректирующими способностями будем строить на следующих простых положениях и оценках:

а) избыточные символы вычисляются на основе соотношения

$$\boldsymbol{H}_{r\times n} \times (\dot{m}_N)^{\mathrm{T}} = 0, \qquad (5)$$

здесь «т» означает транспонирование;

б) в системе применяется синдромный метод декодирования извлеченного из s' кодового слова $\dot{m}_N'$;

в) синдром ошибки Sr – r-разрядный вектор, вес Хемминга которого равен нулю при отсутствии ошибок в $\dot{m}_N'$ ($\dot{m}_N' = \dot{m}_N$); при этом в общем случае

$$Sr = \boldsymbol{H}_{r\times n} \times (\dot{m}_N')^{\mathrm{T}}, \qquad (6)$$

и в соответствии с (5), (6) Sr равен сумме (mod 2) тех вектор-столбцов $\boldsymbol{H}_{r\times n}$, позиции которых соответствуют местоположению ошибок в $\dot{m}_N'$;

г) в извлеченном ($M'\times f_{\mathrm{ECC}}$) кодовом слове $\dot{m}_N'$ длиной n битов могут появиться n одиночных ошибок (в любом отдельно взятом символе) и $\lceil n/2 \rceil$ двойных парных (смежных) ошибок: в битах 1-2, 3-4, 5-6 и т. д.;

д) последнее положение – с учетом п. б) – означает, что конструкция матрицы **A** должна быть такой, чтобы суммы по модулю два смежных парных вектор-стобцов были разными и имели вес Хемминга больший 1(например, $wt(\mathbf{a}_1+\mathbf{a}_2) \geq 1 \bmod 2$ или $wt(\mathbf{a}_3+\mathbf{a}_4) \geq 1 \bmod 2$ и т.д.);

е) с учетом сформулированных положений должно выполняться следующее условие:

$$2^r \geq 1 + n + n/2$$

или иначе

$$2^r \geq 1 + (l+r) + (l+r)/2. \qquad (7)$$

Решение неравенства (5) относительно r дает такой результат:

$$r \geq \log_2 l + 2. \qquad (8)$$

Последнее выражение характеризует разрабатываемую конструкцию кода как приближение по параметру относительной избыточности (R) к коду Хемминга с минимальным кодовым расстоянием d=4.

2. Результаты и их обсуждение.

В статье представлен новый механизм использования избыточного кода совместно со стеганографическим преобразованием информации на основе, например, метода LSB. Основная особенность этого механизма заключается не в адаптации закодированной тайной информации к содержанию контейнера, а в защите ее целостности. Предложенная конструкции кода для обнаружения и коррекции одиночных и двойных парных (в смежных символах) ошибок характеризуется меньшей избыточностью по сравнению с кодами для коррекции одиночных и двойных независимых ошибок. Классическим примером последних являются коды БЧХ [9, с.87–95, с.566]. Для сравнения в таблице приведены соответствующие характеристики двух кодов. В таблице используется следующий формат записи данных: БЧХ | ПК. Видно, что предложенная конструкция кода по количеству избыточных символов r для фиксированного l обладает преимуществом перед БЧХ.

Таблица – Сравнение параметров предложенного кода (ПК) и кода БЧХ

l	4	8	16
r	6 | 4	8 | 5	10 | 6
n	10 |8	16 | 13	26 | 22
R	2,5 | 2,0	2,0 | 1,625	1,625 | 1,373

Ниже также приведены примеры проверочных матриц предложенных кодов (8, 4): для $n = 4$ $r = 4$ и (13, 5) – $n = 13$ $r = 5$. В первом случае информационное слово состоит из полубайта, во втором – из 8 битов, что соответствует кодированию одного символа текста в кодах ASCII:

$$\boldsymbol{H}_{4\times 8} = \begin{matrix} 1101 & 1 & \\ 0111 & & 1 \\ 1111 & & \\ 0110 & & \end{matrix}, \qquad \boldsymbol{H}_{5\times 13} = \begin{matrix} 11011100 & 1 & \\ 11110110 & & 1 \\ 01101111 & & \\ 11110001 & & \\ 01001011 & & \end{matrix}.$$

Отмечаем, что если пары столбцов считать слева направо, то одиночным и парным ошибкам соответствуют не повторяющиеся синдромы.

Библиографические ссылки

1. Urbanovich P., Shutko N. Theoretical Model of a Multi-Key Steganography System // Recent Developments in Mathematics and Informatics. Contemporary Mathematics and Computer Science. Vol. 2, Chapter 11. Lublin: KUL, 2016. P. 181–202.
2. Crandall R. Some notes on steganography. Posted on steganography mailing list. 1998. URL: http://os.inf.tu-dresden.de/~westfeld/crandall.pdf .
3. Westfeld A. F5: a steganographic algorithm // Proc. 4th Int. Workshop Information Hiding 2001, Lecture Notes in Computer Science, vol. 2137. 2001. P. 289–302.

4. Fridrich J., Goljan M. and Soukal D. Efficient wet paper codes. In Proceedings of Information Hiding, Springer Verlag, 2005.
5. Zhang, W., Zhang, X., Wang, S. Maximizing Steganographic Embedding Efficiency by Combining Hamming Codes and Wet Paper Codes. In: Solanki, K., Sullivan, K., Madhow, U. (eds) Information Hiding. IH 2008. Lecture Notes in Computer Science, vol 5284. Springer, Berlin, Heidelberg, 2008. https://doi.org/10.1007/978-3-540-88961-8_5.
6. Filler T., Judas J., Fridrich J. Minimizing additive distortion in steganography using syndrome-trellis codes // IEEE Transactions on Information Forensics and Security. 2011 № 6(3-2). P. 920–935.
7. Li W., Zhang W., Li L., Zhou H., Yu N. Designing near-optimal steganographic codes in practice based on polar codes // IEEE Transactions on Communications. 2020. № 68(7). P. 3948–3962.
8. Урбанович П.П., Алексеев В.Ф., Верниковский Е.А. Избыточность в полупроводниковых интегральных микросхемах памяти. Мн.: Навука і тэхніка, 1995. 262 с.
9. Мак-Вильмс Ф.Дж., Слоэн Н.Дж. Теория кодов, исправляющих ошибки: Пер. с англ. М.: Связь, 1979. 744 с.

УДК [004.056+003.26](075.8)

ИСПОЛЬЗОВАНИЕ ОСОБЕННОСТЕЙ ФОРМАТА XML В МЕТОДАХ ТЕКСТОВОЙ СТЕГАНОГРАФИИ

П.П. Урбанович, О.А. Нистюк, М.Г. Савельева, Н.П. Шутько, А.Н. Николайчук

Белорусский государственный технологический университет,
ул. Свердлова, 13а, 220005 Минск, Беларусь,
для корреспонденции: shutko_bstu@mail.ru

Описаны новые методы текстовой стеганографии, основанные на модификации пространственно-геометрических и цветовых параметров элементов электронных текстов-контейнеров и учете особенностей формата XML. Оценены пропускная способность и стойкость методов к модификации стеганоконтейнера.

Ключевые слова: XML-формат; текстовая стеганография; LSB-методы; модель RGB; стеганографическая стойкость.

USING OF THE XML FORMAT FEATURES IN THE METHODS OF TEXT STEGANOGRAPHY

P.P. Urbanovich, O.A. Nistyuk, M.G. Saveleva, N.P. Shutko, A.N. Nikolaichuk

Belarusian State Technological University,
Sverdlova str., 220005 Minsk, Belarus,
corresponding author: shutko_bstu@mail.ru

New methods of text steganography, based on the modification of the spatial-geometric and color parameters of the elements of electronic text-containers and taking into account the features of the XML format are described. The covert channel capacity and resistance of the methods to the steganocontainer modification are characterized.

Keywords: XML format; text steganography; LSB methods; RGB model; steganographic resistance.

Введение

Как известно, передача и защита информации на основе стеганографии основана на сохранении в тайне самого факта реализации стеганографического преобразования. Это обстоятельство влияет на две взаимосвязанные цели исследований в данной предметной области: стеганографические методы должны обеспечивать высокую пропускную способность стеганоканала при максимально высоком уровне скрытности [1].

Указанные цели и соответствующие им направления разработки прикладных стеганометодов, как правило, основаны на такой модификации параметров стеганоконтейнера (при размещении тайной информации), которая сводила бы на нет эффективность визуальной атаки – с одной стороны, и не влияла бы на целостность осажденной (передаваемой) информации при случайной или преднамеренной модификации стеганоконтейнера. Если стеганоканал создается на основе электронных текстовых документов, то модифицировать можно как отдельные пространственно-цветовые параметры текста [2–5], так и отдельные атрибуты текстового файла-контейнера [6]. При этом уровень скрытности осаждаемого в контейнер сообщения, что нами отождествляется со скрытностью стеганоканала, связан с относительной частью модифицируемого параметра – по аналогии с относительной частью наименее значащих битов (least significant bits, LSB), используемых при размещении тайного сообщения, по отношению к битовой длине используемого параметра. Последний, например, при цветовой кодировке пикселей в одном цветом канале модели RGB составляет 8 битов.

В [7, 8] описаны общие концепции и основные особенности новых методов тестовой стеганографии, развивающих и дополняющих теорию и практику стеганографических преобразований текстовых документов-контейнеров на основе формата XML. В данной статье представлены новые результаты, характеризующие методы из [7, 8].

1. Основная часть

Использование XML-формата в текстовой стеганографии. Основополагающая идея использования пространственно-геометрических и цветовых параметров элементов текстовых документов в качестве носителей тайной информации базируется на специфике формата XML [5]. С его помощью решаются, в частности, задачи хранения и транспортировки данных в процессоре MS Word при обработке как векторной, так и растровой графики (текстовый документ также можно рассматривать как графический объект).

Файл формата *DOCX* представляет собой ZIP-архив, который содержит два типа файлов: файлы XML с расширениями *xml* и *rels* и медиафайлы (например, изображения). Можно сказать, что *DOCX*-файл представляет собой набор сжатых файлов формата XML, причем все текстовое содержимое электронного документа MS Word формата *DOCX* находится в одном файле – *document.xml*.

Для описания особенностей форматирования текста используются, как и в других языках разметки, теги. Например, тег описания свойств абзаца (<*w:pPr*>) содержит в себе вложенный тег описания межстрочного интервала, например, <*w:spacing w:lineRule="exact" w:line="360"/*>, который обозначает, что высота межстрочного интервала задана точно и составляет 18 пунктов (пт). Для описания форматирования отдельных символов используется тег <*w:rPr*>. Например, в конструкции <*w:rPr*><*w:sz w:val="28"/*> <*w:szCs w:val="28"/*></*w:rPr*> параметр <*w:sz w:val="28"/*> измеряется в ½ пт и в данном случае указывает, что кегль текста равен 14 пт.

Как следует из изложенного, размещение тайной информации в электронном текстовом документе может осуществляться путем модификации определенных параметров, отвечающих за форматирование текста. Далее рассмотрим важные особенности предлагаемых стеганометодов.

Метод на основе растровой графики и цветовой модели RGB. В качестве базового элемента контейнера, цветовые параметры которого модифицируются в модели *RGB* при размещении тайной информации, выступает пиксель изображения. Внедрение/извлечение информации происходит в пикселях, имеющих одинаковое значение (одно из 256) в одном или нескольких цветовых каналах (R, G, B).

Для внедрения сообщения *M* в контейнер *C* необходимо выбрать массив пикселей, для которых совпадает значение координат одного или двух цветовых каналов. Пояснение к этому дает рисунок 1, на котором в увеличении представлен фрагмент буквы. В изображениях с большим количеством полутонов, монохроматических или черно-белых изображениях выбор пикселей, в которых будет происходить внедрение, целесообразно осуществлять по двум цветовым каналам. При этом непосредственно для внедрения информации в выбранные пиксели используется один канал.

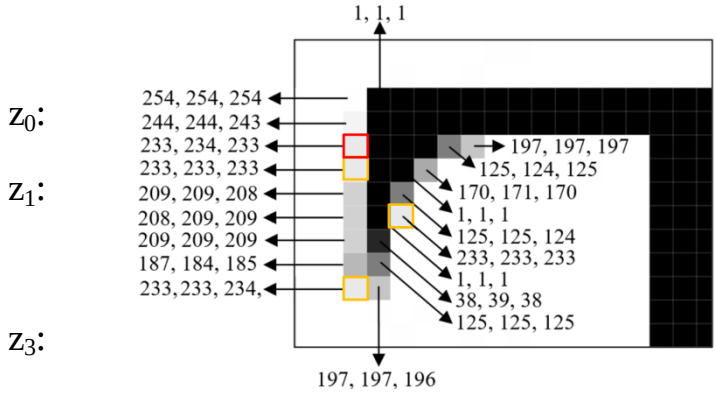


Рисунок 1 – Пояснение к алгоритму анализа и выбора пикселей массива *Z*

Важным шагом алгоритма внедрения является выбор массива пикселей, Z ($Z=\{z_i\}$, здесь $i = \overline{0,\text{length}(Z)}$). Необходимо также определить следующие элементы: c_{RGB} – цветовой канал с совпадающими цветовыми параметрами пикселя, $c_{RGB} \in R, G, B$, c_{RGB}' – цветовой канал для внедрения сообщения M, $c_{RGB}' \in R, G, B$, s_{jn} – пиксельный элемент документа C, $s_{jn} \in C$ ($C=\{s_{jn}\}$, $j = \overline{0,t}$, $n = \overline{0,r}$), t и r – размер C: соответственно ширина и высота в пикселях, φ – ключевое значение цветового кода канала c_{RGB}, $\varphi \in \{0, 1, \ldots, 255\}$. Последний параметр используется для увеличения пропускной способности создаваемого скрытого канала. Для этого следует провести анализ того, в каком цветовом канале имеется больше пикселей с одинаковым значением цветового кода (c_{RGB}) и выбрать это значение в качестве параметра φ. Канал для внедрениям (c_{RGB}') выбирается произвольно из оставшихся двух (или оба). Канал c_{RGB}' не должен использоваться при формировании массива пикселей Z. Например, после проанализированного фрагмента изображения-контейнера (часть буквы; рисунок 1) можно сказать, что $c_{RGB} = R$, $c_{RGB}' = G$ (как следующий после R), $\varphi = 233$.

Из массива Z выбирается базовый пиксель. Внедрение M будет происходить в канал c_{RGB}' при сравнении значений цветовых кодов канала c_{RGB}' пикселя для внедрения, $c_{RGB}'(z_0)$ ($z_i \in Z$) и базового пикселя, $c_{RGB}'(z_i)$. В этом примере базовый пиксель имеет цветовой код (233, 234, 233). Далее внедрение будет происходить при сравнении значений кода канала G базового (первого) и второго, базового и третьего и т. д.; в конечном итоге – базового и n-ного пикселей массива Z.

Для извлечения внедренного сообщения необходимо выбрать массив пикселей Z_D (пример показан на рисунке 2), где совпадает значение кода одного или нескольких цветовых каналов (по аналогии с формированием Z).

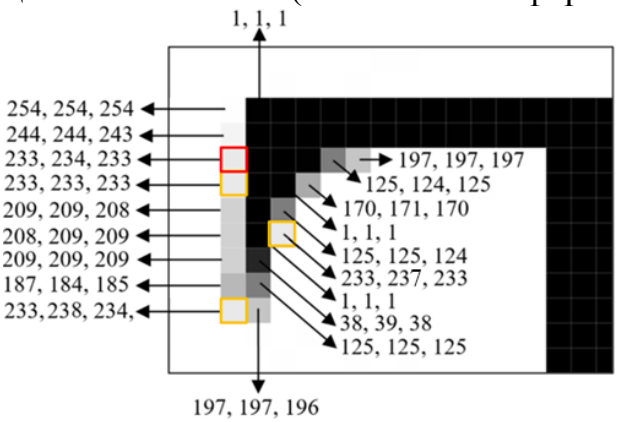


Рисунок 2 – Пример массива Z_D на фрагменте изображения

В отличие от массива Z в массив Z_D помещаются пиксели, для которых $c_{RGB}'(s_{jn})$ отличается от $c_{RGB}'(z_0)$ на $2Q$ единиц в любую сторону (диапа-

зон для выбора: $c_{RGB}'(z_0) - 2Q \le c_{RGB}'(s_{jn}) \le c_{RGB}'(z_0) + 2Q,)$. Параметр Q выбирается из условия: $4 < Q < 10$; чем шире граница указанного диапазона, тем выше пропускная способность канала, а указанные нижний и верхний пределы обеспечивают высокую устойчивость канала к визуальным атакам, что установлено многочисленными тестами. В качестве ключей стеганографического преобразования [4] может использоваться информация о том, какой канал (или несколько каналов) используется для выбора пикселей для внедрения *M*, координаты базового пикселя в массиве *Z* или алгоритм выбора базового пикселя из массива пикселей для внедрения, канал для внедрения, значение φ, значение *Q*.

Для оценки эффективности метода по параметрам пропускной способности и устойчивости к стеганоконтейнера к модификациям разработано программное приложение, с помощью которого, в частности, установлено, что текстовый документ формата *PNG* с тайной информацией сохраняет целостность этой информации после конвертации в форматы *GIF, BMP, TIFF* (со сжатием и без сжатия), однако при конвертации в *JPG* примерно 70-75% битов исходного сообщения *M* являются ошибочными. Пропускная способность канала на основе метода примерно соответствует аналогичному параметру для семейства методов на основе LSB. Важным является то, что изображения с большим количеством полутонов, черно-белые и монохромные изображения будут обеспечивать сравнительно более высокую пропускную способность, так как они построены на основе большего количество пикселей с совпадающими кодами цветового канала.

Метод на основе модификации параметров контура символов текста является близким аналогом методов, основанных на модификации цвета символов текста, а также параметров апроша и кернинга [2, 3]. Параметры контура можно легко найти, выбрав пункт меню *Главная* в среде MS Word. К основным из таких параметров относятся: цвет, прозрачность, ширина, составной тип, тип штриха и др. Глубина изменения каждого из параметров влияет на пропускную способность и устойчивость преобразования к визуальным атакам в известном соотношении (лучше одно – хуже другое). Для примера на рисунке 3 показан вид букв с контуром и без контура. Даже при значительном увеличении контур остается визуально незаметным.

без контура

с контуром

Рисунок 3 – Примеры шрифтового оформления с контуром и без него

При реализации метода необходимо принять во внимание символы, которые не могут быть дополнены контуром. К ним относятся: ", #, $, %, &, ', (,), *, +, «,», -, ., /, :, ;, <, =, >, ?, @, [,], ^, _, `, {, |, }, ~, –, \\s, \.

Для анализа эффективности метода авторами создано отдельное приложение. Оригинальный текст с внедренным сообщением конвертировался в форматы *PDF, TXT, DOC* и обратно. После обратной конвертации сообщение *M* восстановить не удалось. Однако использование всех вышеперечисленных архиваторов не влияет на целостность *M* после распаковки архива.

Выводы

Для оценки эффективности (пропускной способности, устойчивости к случайным или преднамеренным модификациям стеганоконтейнера с размещенной тайной информацией) разработаны специальные программные средства, зарегистрированной в Государственном реестре информационных ресурсов РБ. Установлено, что метод на основе растровой графики и цветовой модели RGB обеспечивает целостность осажденной информации при конвертации стеганоконтейнера в большинство основных форматов, а метод на основе модификации параметров контура символов текста – при конвертации в форматы *PDF, TXT, DOC* и обратно. Пропускная способность канала на основе предложенных методов примерно соответствует аналогичному параметру для семейства методов на основе LSB.

Библиографические ссылки

1. Subramanian N., Elharrouss O., Al-Maadeed S., Bouridane A. Image Steganography: A Review of the Recent Advances // IEEE Access. 2021. № 9. P. 23409 –23423. DOI: 10.1109/ACCESS.2021.3053998.
2. Shutko N., Urbanovich P., Zukowski P. A method of syntactic text steganography based on modification of the document-container aprosh // Przegląd Elektrotechniczny. 2018. № 94(6). P. 82-85. DOI:10.15199/48.2018.06.15.
3. Шутько Н.П. Алгоритмы реализации методов текстовой стеганографии на основе модификации пространственно-геометрических и цветовых параметров текста // Труды БГТУ. 2016. № 6. С. 160–165.
4. Urbanovich P., Shutko N. Theoretical Model of a Multi-Key Steganography System // Recent Developments in Mathematics and Informatics. Contemporary Mathematics and Computer Science. Vol. 2, Chapter 11. Lublin: KUL, 2016. P. 181–202.
5. Блинова Е.А., Сущеня А.А. Применение нескольких стеганографических методов для осаждения скрытых данных в электронных текстовых документах // Системный анализ и прикладная информатика. 2019. № 2. С. 32–38. URL: https://doi.org/10.21122/2309-4923-2019-2-32-38
6. Урбанович П.П., Юрашевич Д.Э. Использование системных свойств и парамет-

ров текстовых файлов в стеганографических приложениях // Теоретическая и прикладная криптография: материалы междунар. научной конференции. Минск, 20–21 октября 2020 г. Минск: БГУ, 2020. С. 68–73.
7. Нистюк О.А. Защита текстовой информации с помощью добавления контура к символам текста // Информационные технологии: материалы 86-й научно-техн. конф. профессорско-препод. состава, научных сотр. и аспирантов, Минск, 31.01 – 12.02 2022 г. Минск: БГТУ, 2022. С. 59–62.
8. Савельева М.Г. Метод стеганографического внедрения тайной информации в WEB-документы на основе растровой графики // Информационные технологии: материалы 86-й научно-техн. конф. профессорско-препод. состава, научных сотр. и аспирантов, Минск, 31.01 – 12.02 2022 г. Минск: БГТУ, 2022. С. 52–54.

УДК 519.671:003.26

ЗАЩИТА ИНФОРМАЦИИ И СТОХАСТИКА

Ю.С. Харин[1, 2]

[1]*Научно-исследовательский институт прикладных проблем математики и информатики,*
[2]*Белорусский государственный университет,*
пр. Независимости, 4, 220030, г. Минск, Беларусь, kharin@bsu.by

Рассматривается проблема статистического тестирования дискретно-значных временных рядов на «чистую случайность». Предложены модели отклонений от «чистой случайности», методы и алгоритмы обнаружения этих отклонений. Приводятся результаты компьютерных экспериментов.

Ключевые слова: временной ряд; цепь Маркова; статистический тест.

INFORMATION PROTECTION AND STOCHASTICS

Yu.S. Kharin[a,b]

[a]*Research Institute for Applied Problems of Mathematics and Informatics,*
[b]*Belarusian State University, 4 Nezavisimosti avenue, Minsk 220030, Belarus,*
kharin@bsu.by

Problem of statistical testing for “pure randomness” of discrete-valued time series is considered. Models of deviations from “pure randomness”, methods and algorithms for detection of these deviations are proposed. Results of computer experiments are given.

Keywords: time series; Markov chain; statistical test.

Введение

В современных системах обеспечения информационной и компьютерной безопасности важнейшим способом защиты информации является криптографический способ, позволяющий с гарантированной стойкостью решить главные практические задачи: 1) конфиденциальность; 2) аутентификация источника сообщения; 3) проверка целостности; 4) невозможность отречения от авторства. Криптографический способ базируется на новой науке Криптологии [1], объединяющей Криптографию и Криптоанализ. Криптология и Стохастика тесно связаны: Стохастика представляет математический инструментарий для решения задач Криптологии, Криптология стимулирует Стохастику к разработке новых моделей для исследования сложных последовательностей, циркулирующих в криптосистемах.

Настоящий доклад посвящен представлению и использованию новых стохастических моделей в криптографической защите информации.

1. Проблема «чистой случайности» в защите информации

Многие задачи криптологии (статистическое тестирование криптографических генераторов, статистический криптоанализ, разностный криптоанализ, линейный криптоанализ, криптоатаки по побочным каналам, стеганография) сводятся к задаче различения некоторой зарегистрированной последовательности символов $x_1, x_2, \ldots$ от «чисто случайной» и оценки величины этого различия.

Математической моделью последовательностей, порождаемых генераторами, а также последовательностей, возникающих в различных узлах средств криптографической защиты информации, является дискретный временной ряд (ДВР). ДВР – это случайный процесс $x_t \in A$ на вероятностном пространстве $(\Omega, F, \mathbf{P})$ с дискретным временем $t \in \mathbf{N} = \{1, 2, \ldots\}$ и дискретным множеством состояний (алфавитом) $A = \{0, 1, \ldots, N-1\}$ мощности $|A| = N$, $2 \le N < +\infty$.

В криптологии [1] согласно Шенноновской теории совершенных криптосистем большое внимание уделяется так называемому «чисто случайному» ДВР – равномерно распределенной случайной последовательности (РРСП) $x_1, x_2, \ldots \in A$, обладающей двумя свойствами:

C_1) для любого числа $n \in \mathbf{N}$ и произвольных индексов $1 < t_1 < \ldots < t_n$ случайные элементы $x_{t_1}, \ldots, x_{t_n}$ независимы в совокупности;

C_2) для любого $t \in \mathbf{N}$ случайная величина x_t имеет равномерное на A распределение вероятностей: $\mathbf{P}\{x_t = i\} = N^{-1}$, $i \in A$.

В настоящее время известно более сотни методов генерации последовательностей, по своим свойствам приближающихся к РРСП. Еще больше разработано статистических тестов криптографических генераторов, заключающихся в проверке простой гипотезы $H_0 = \{\{x_t\}$ есть РРСП$\}$ против сложной альтернативы $H_1 = \overline{H}_0 = \{$нарушены свойства $C_1 \cup C_2\}$. Обзор статистических тестов показывает:

1) многие из существующих тестов не ориентированы на проверку главного свойства C_1, а лишь частных случаев свойств C_1, C_2, т.е. частных случаев альтернативы H_1;

2) многие тесты построены «эвристически» и не фиксируют H_1;

3) многие тесты не имеют оценок мощности;

4) при включении нескольких тестов в батарею не удается оптимизировать «составной» тест.

В связи с этим актуальна проблема разработки адекватных стохастических моделей отклонений H_1 от модели РРСП и построения тестов для обнаружения и оценивания таких отклонений.

2. Модели ДВР на основе семейства отклонений от s-мерной равномерности и их энтропийное тестирование

Определим вложенное в H_1 семейство «альтернатив s-мерной неравномерности»: $H_{1(s)} = \{\{x_1, x_2, \ldots\} = \{X_1, X_2, \ldots\}\} \subset H_1$, где $X_1, X_2, \ldots \in A^s$ – независимые одинаково распределенные s-фрагменты (слова) над алфавитом A с некоторым s-мерным дискретным распределением вероятностей $\mathbf{P}_{i_1,\ldots,i_s} = \mathbf{P}\{x_1 = i_1, \ldots, x_s = i_s\}$, $i_1, \ldots, i_s \in A$, отличным от равномерного: $\Delta_s = \sum_{i_1,\ldots,i_s \in A} \left| \mathbf{P}_{i_1,\ldots,i_s} - N^{-s} \right| > 0$, $\sum_{i_1,\ldots,i_s \in A} \mathbf{P}_{i_1,\ldots,i_s} \equiv 1$. Это семейство моделей ДВР обладает двумя свойствами: 1) при $s \to \infty$ семейство этих альтернатив имеет в пределе альтернативу $H_1 = \bar{H}_0$ общего вида; 2) чем меньше Δ_s, тем ближе альтернатива $H_{1(s)}$ к H_0.

Обозначим: $\{x_1, x_2, \ldots, x_T\} = \{X_1, X_2, \ldots, X_M\}$ – наблюдаемая реализация выходной последовательности генератора длиной $T = M \cdot s$, разбитая на M непересекающихся фрагментов длины s; $I\{B\}$ – индикатор события B; статистическая оценка для $\mathbf{P}_{i_1,\ldots,i_s}$

$$\hat{\mathbf{P}}_{i_1,\ldots,i_s} = \frac{1}{M} \sum_{m=1}^{M} I\{X_m = (i_1, \ldots, i_s)\}, \; i_1, \ldots, i_s \in A. \qquad (1)$$

Тест обобщенного отношения правдоподобия для проверки H_0, $H_{1(s)}$ на основе статистик (1) имеет вид:

$$\text{принимается} \begin{cases} H_0, \text{ если } \hat{H}_s - s \ln N > -\dfrac{1}{2M} G^{-1}_{N^s - 1}(1 - \varepsilon), \\ H_{1(s)} \text{ в противном случае,} \end{cases} \qquad (2)$$

$\hat{H}_s = -\sum_{i_1,\ldots,i_s \in A} \hat{P}_{i_1,\ldots,i_s} \ln \hat{P}_{i_1,\ldots,i_s}$ – статистическая оценка s-мерной энтропии Шеннона, $G_K^{-1}(\cdot)$ – обратная функция распределения хи-квадрат с K степенями свободы, $\varepsilon \in (0,1)$ – заданный уровень значимости теста.

Тест (1), (2) мы предлагаем использовать для визуализации процесса принятия решений в виде так называемого «энтропийного профиля (портрета)» – графика зависимости нормированного отклонения оценки s-мерной энтропии от ее математического ожидания при H_0 (см. рис. 1, 2 для $N = 2,\ \varepsilon = 0.05,$ где штриховые линии – границы области решений):

$$\alpha(s) = 2M\left(\hat{H}_s - s\ln N\right) / G_{N^s-1}^{-1}(1-\varepsilon),\quad s \in \{s_{min}, s_{min}+1, \ldots, s_{max}\}\ . \qquad (3)$$

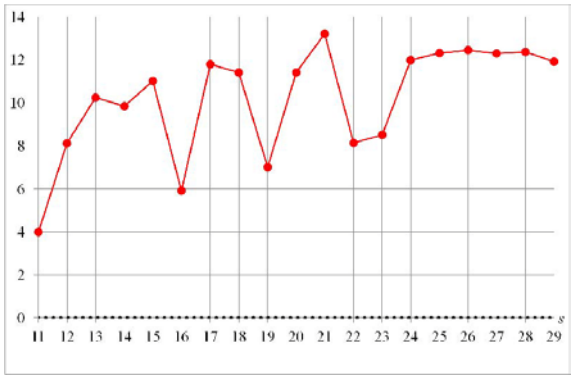


Рисунок 1 – Энтропийный профиль $\ln|\alpha(s)|$ нелинейного регистра сдвига порядка 24 ($T = 2^{32}/s$)

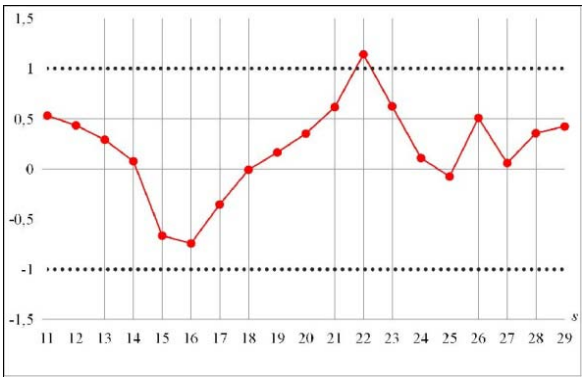


Рисунок 2 – Энтропийный профиль $\alpha(s)$ генератора BelT (СТБ 34.101.27-2011, $T = 2^{29}/s$)

Отметим еще, что вместо энтропии Шеннона в (1) – (3) могут использоваться энтропийные функционалы Реньи и Тсаллиса [2].

3. Модели ДВР на основе марковских зависимостей высокого порядка и их статистическое тестирование

3.1. Тестирование на основе цепи Маркова высокого порядка

Учитывая, что универсальной моделью стохастической зависимости элементов выходной последовательности $\{x_t\}$ криптографического генератора является цепь Маркова достаточно высокого порядка s, определим вложенное в $H_1 = \overline{H}_0$ семейство альтернатив марковской зависимо-

сти: $H_1^{(s)} = \{\{x_t\}$ – однородная цепь Маркова порядка s с $(s+1)$-матрицей переходов $\mathbf{P}\}$, где $\mathbf{P} = \left(p_{i_1,\ldots,i_{s+1}}\right)$, $i_1,\ldots,i_{s+1} \in A$,

$$p_{i_1,\ldots,i_{s+1}} = \mathbf{P}\{x_{t+1} = i_{s+1} \mid x_t = i_s,\ldots,x_{t-s+1} = i_1\}, \Delta_s = \sum_{i_1,\ldots,i_S \in A} \left| p_{i_1,\ldots,i_{s+1}} - N^{-1} \right| > 0. \qquad (4)$$

Тест обобщенного отношения правдоподобия для проверки гипотез $H_0, H_1^{(s)}$ основан на оценке $\hat{h}_s$ условной энтропии $h_s = H\{x_t \mid x_{t-1},\ldots,x_{t-s}\}$:

$$\text{принимается} \begin{cases} H_0, \text{если } \hat{h}_s - \ln N > -G_f^{-1}(1-\varepsilon)/\left(2(T-s)\right),\ f = N^s(N-1), \\ H_1^{(s)} \text{ в противном случае.} \end{cases} \qquad (5)$$

Аналогично (3) с помощью $\hat{h}_s$ строится энтропийный профиль.

Тесты (2), (5), анализирующие стохастические зависимости глубины s в выходной последовательности $\{x_t\}$, требуют экспоненциально растущей с ростом порядка s длины анализируемой последовательности $T = O\left(N^{s+1}\right)$. Для преодоления этой трудности целесообразно использовать «малопараметрические модели цепей Маркова высокого порядка» [1, 3], т.е. модели цепей Маркова s-го порядка, для которых $\left(N^s \times N\right)$-матрица вероятностей переходов зависит от «малого» числа параметров $D << N^s(N-1)$; $\kappa = D/\left(N^s(N-1)\right) << 1$ – коэффициент сжатия, равный относительному числу параметров модели.

3.2. Построение малопараметрических цепей Маркова

Подход I: «сжатие множества значений элементов матрицы» $\mathbf{P}$.

Пусть $Q = \left(q_{j_i,\ldots,j_r,j_{r+1}}\right)$ – некоторая $(r+1)$-мерная матрица, $1 \le r < s$, $\sum_{j_{r+1} \in A} q_{j_i,\ldots,j_r,j_{r+1}} \equiv 1, 0 \le q_{j_i,\ldots,j_r,j_{r+1}} \le 1$; $B(\cdot)\colon A^s \to A^r$ – некоторая дискретная функция. С помощью $B(\cdot)$ $(s+1)$-мерная матрица $\mathbf{P}$ «сжимается» в $(r+1)$-мерную матрицу Q:

$$p_{i_1,\ldots,i_s,i_{s+1}} = q_{B(i_1,\ldots,i_s),i_{s+1}};\ \kappa_{\mathrm{I}} = N^{r-s} \le 1. \qquad (6)$$

Примеры малопараметрических ДВР: $\mathrm{MC}(s,r)$, $\mathrm{MCCO}(s,L)$, VLMC [4].

Подход II. Этот подход заключается в использовании порождающего уравнения для условного распределения вероятностей (4) будущего состояния $x_t \in A$ при условии предыстории $X_{t-s}^{t-1} = \left(x_{t-1},\ldots,x_{t-s}\right)' \in A^s$:

$$p_{i_1,\ldots,i_s,i_{s+1}} = q_{i_{s+1}}\left(\theta\left(i_1,\ldots,i_s;a\right)\right),\ i_1,\ldots,i_{s+1} \in A, \qquad (7)$$

где $\left\{q_j(\theta): j \in A\right\}$ – некоторое стандартное вероятностное распределение на A, зависящее от параметра $\theta = \left(\theta_j\right) \in \Theta \subseteq R^L$; $\theta = \theta\left(i_1,\ldots,i_s;a\right)$ – некоторая функция, известная с точностью до вектора параметров $a = \left(a_k\right) \in R^m$. Коэффициент сжатия: $\kappa_{\text{II}} = m/\left(N^s\left(N-1\right)\right) \le 1$.

Примеры малопараметрических ДВР: модель Джекобса – Льюиса, MTD-модель, $\text{DAR}(s)$, $\text{BCNAR}(s)$, $\text{BiCNAR}(s)$, $\text{PCNAR}(s)$.

4. Малопараметрические модели ДВР на основе подхода I и их статистическое тестирование

4.1. Цепь Маркова $\text{MC}(s,r)$ порядка s с r частичными связями

Эта модель определяется (6) с $B(j_1\ldots,j_s) = \left(j_{m_1^0},\ldots,j_{m_r^0}\right)$ [1, 3]:

$$p_{J_1^{s+1}} = p_{j_1,\ldots,j_s,j_{s+1}} = q_{j_{m_1^0},\ldots,j_{m_r^0},j_{s+1}},\ J_1^{s+1} \in A^{s+1}, \qquad (8)$$

где $J_i^k = \left(j_i, j_{i+1},\ldots,j_k\right) \in A^{k-i+1}$ – последовательность $k-i+1$ индексов $\left(k \ge i\right)$; r – число связей; $M_r^0 = \left(m_1^0,\ldots,m_r^0\right)$ – вектор с r упорядоченными компонентами $1 = m_1^0 < m_2^0 < \ldots < m_r^0 \le s$, называемый шаблоном связей; $Q = \left(q_{J_1^{r+1}}\right)_{J_1^{r+1} \in A^{r+1}}$ – $\left(r+1\right)$-мерная стохастическая матрица.

Статистическую оценку $\hat{Q}$ удобно использовать для визуализации отклонения от гипотезы H_0 (для которой $q_{i_1,\ldots,i_{r+1}} = N^{-1}$). На рис. 3, 4 представлены результаты такой визуализации для генератора со случайной обратной связью и генератора BelT (СТБ 34.101.27-2011 в режиме гаммирования) соответственно; здесь красный цвет – оценка условной вероятности перехода в «0» $\hat{q}_{K_1^r,0}$, зеленый – в «1» $\hat{q}_{K_1^r,1}$); по оси абсцисс откладывается $K_1^r = B\left(J_1^s;\hat{M}_r\right) \in A^r$.

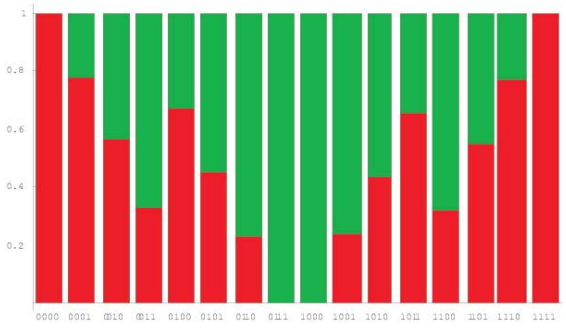


Рисунок 3 – Оценка $\hat{Q}$
$\left(s=64,\ r=4,\ T=10^5\right)$

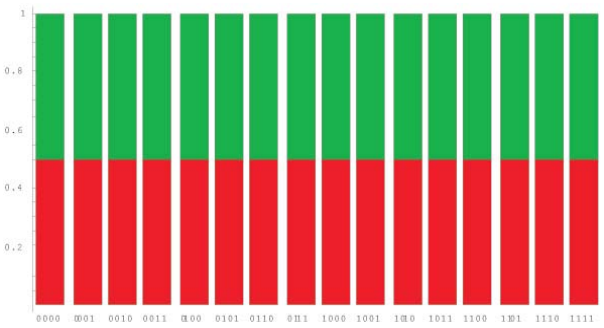


Рисунок 4 – Оценка $\hat{Q}$
$\left(s=32,\ r=4,\ T=8\cdot10^6\right)$

4.2. Модель Джекобса – Льюиса

Эта модель порождается стохастическим разностным уравнением [5]:

$$x_t=\mu_t x_{t-\eta_t}+\left(1-\mu_t\right)\xi_t, \tag{9}$$

где $t>s$, $\{\xi_t,\eta_t,\mu_t\}$ – независимые в совокупности случайные величины с вероятностными распределениями:

$$\mathbf{P}\{\mu_t=1\}=1-\mathbf{P}\{\mu_t=0\}=\rho;\ \mathbf{P}\{\xi_t=k\}=\pi_k\ ,\ k\in A\ ,\ \sum_{k\in A}\pi_k=1\ ;$$

$$\mathbf{P}\{\eta_t=i\}=\lambda_i,\ i\in\{1,2,\ldots,s\},\ \sum_{i=1}^{s}\lambda_i=1,\ \lambda_s\neq0; \tag{10}$$

$$\mathbf{P}\{x_1=k\}=\ldots=\mathbf{P}\{x_s=k\}=\pi_k,\ k\in A.$$

Число параметров этой модели (9), (10) линейно зависит от s, так что коэффициент сжатия $\kappa_{JL}=\left(N+s-1\right)/\left(N^s\left(N-1\right)\right)$. Методы и алгоритмы статистического анализа этой модели представлены в [1].

4.3. MTD-модель Рафтери

MTD (Mixture Transition Distribution)-модель [6] определяется следующим частным случаем уравнения (7): $p_{i_1,\ldots,i_s,i_{s+1}}=\sum_{j=1}^{s}\lambda_j q_{i_j,i_{s+1}},i_1,\ldots,i_{s+1}\in A,$ где $Q=\left(q_{i,k}\right)$ – некоторая стохастическая $\left(N\times N\right)$-матрица, $0\le q_{i,k}\le1$, $\sum_{k\in A}q_{i,k}\equiv1$, $i,k\in A$, $\lambda=\left(\lambda_1,\ldots,\lambda_s\right)'$ – некоторое дискретное распределение вероятностей, $\lambda_1>0$.

Обобщенная MTDg (generalized MTD)-модель определяется параметризацией $(s+1)$-мерной матрицы $\mathbf{P}$: $p_{i_1,\ldots,i_s,i_{s+1}} = \sum_{j=1}^{s} \lambda_j q^{(j)}_{i_j,i_{s+1}}$, $i_1,\ldots,i_{s+1} \in A$, где $Q^{(j)} = \left(q^{(j)}_{i,k}\right)$ – некоторая стохастическая матрица для j-го лага, $\kappa_{\mathrm{MTDg}} = \left(s\left(N(N-1)/2+1\right)-1\right)/\left(N^s(N-1)\right)$. Методы и алгоритмы статистического тестирования даны в [1].

4.4. Биномиальная условно нелинейная авторегрессионная модель BiCNAR(s)

Эта модель порождается биномиальным случаем уравнения (7):

$$p_{i_1,\ldots,i_s,i_{s+1}} = C^{i_{s+1}}_{N-1}\, \theta^{i_{s+1}} (1-\theta)^{N-1-i_{s+1}}, i_{s+1} \in A = \{0,1,\ldots,N-1\},$$

$$\theta = \theta\left(\mathrm{I}^{\mathrm{s}}_1\right) = \mathrm{F}\left(a'\Psi\left(\mathrm{I}^{\mathrm{s}}_1\right)\right),\ \mathrm{I}^{\mathrm{s}}_1 = (i_1,\ldots,i_s)' \in A^s,$$

где $\Psi\left(\mathrm{I}^{\mathrm{s}}_1\right) = \left(\psi_1\left(\mathrm{I}^{\mathrm{s}}_1\right),\ldots,\psi_m\left(\mathrm{I}^{\mathrm{s}}_1\right)\right)'$: $A^s \to R^m$ – вектор-столбец $m \le N^s$ линейно независимых функций, например, полиномов; $\mathrm{F}(\cdot)$: $R^1 \to [0,1]$ – некоторая функция распределения, например, логистическая, нормальная или Коши; $a = (a_1,\ldots,a_m)'$ – вектор-столбец m неизвестных параметров модели. Относительное число параметров модели: $\kappa = m\left(N^s(N-1)\right)^{-1} < 1$.

Методы и алгоритмы статистического анализа BiCNAR(s)-модели, ее частных случаев и обобщений представлены в [7, 8].

Заключение

1. В криптологии актуальна проблема построения и статистического анализа моделей ДВР, адекватно описывающих отклонения от РРСП.
2. Представлены семейства моделей ДВР на основе отклонений от s-мерной равномерности и на основе марковских зависимостей порядка s.
3. Для преодоления «проклятия размерности» представлены подходы к построению малопараметрических цепей Маркова порядка s.
4. Разработаны методы и алгоритмы статистического оценивания параметров и проверка гипотез H_0, H_1 для малопараметрических моделей, построенных на основе предложенных подходов.

5. Теоретические результаты иллюстрируются результатами компьютерных экспериментов по тестированию выходных последовательностей известных криптографических генераторов.

Библиографические ссылки

1. Харин Ю.С., Агиевич С.В., Васильев Д.В., Матвеев Г.В. Криптология. Минск: БГУ, 2014. 512 с.
2. Харин Ю.С., Палуха В.Ю. Энтропийный анализ криптографических генераторов случайных и псевдослучайных последовательностей // Веснік сувязі. 2017. № 146(1). С. 46–49.
3. Харин Ю.С. Цепи Маркова с S-частичными связями и их статистическое оценивание // Доклады НАН Беларуси. 2004. № 48(1). С. 40–44.
4. Buhlmann P., Wyner A.J. Variable length Markov chains // The Annals of Statistics. 1999. № 27(2). P. 480–513.
5. Jacobs P.A., Lewis P.A.W. Discrete time series generated by mixtures I: correlational and runs properties // Journal of the Royal Statistical Society. Ser. B. 1978. № 40(1). P. 94–105.
6. Raftery A. A model for high-order Markov chains // Journal of the Royal Statistical Society. Ser. B. 1985. № 47(3). P. 528–539.
7. Харин Ю.С., Волошко В.А. Биномиальные условно нелинейные авторегрессионные модели дискретных временных рядов и их вероятностные и статистические свойства // Труды Института математики НАН Беларуси. 2019. № 26(1). С. 95–105.
8. Kharin Yu.S., Voloshko V.A., Medved E.A. Statistical estimation of parameters for binary conditionally nonlinear autoregressive time series // Mathematical Methods of Statistics. 2019. № 26(2). P. 103–118.

ИНТЕЛЛЕКТУАЛЬНЫЙ И СТАТИСТИЧЕСКИЙ АНАЛИЗ ДАННЫХ, ПРИНЯТИЕ РЕШЕНИЙ

УДК 519.237:330.43

STATISTICAL ANALYSIS AND ECONOMETRIC MODELING OF THE COVID-19 PANDEMIC

V.I. Malugin, A.K. Kornievich, V.A. Potapovich

Belarusian State University, Nezalezhnosti Ave., 4, 220030, Minsk, Belarus
Malugin@bsu.by

This paper presents the results of solving the following tasks: development of statistical methods for classifying countries in the European Region by the intensity of the COVID-19 pandemic; building country ratings that characterize the intensity of the pandemic; assessment of the relationship of country ratings with the economic indicators of countries; development of econometric models of the epidemic process in the Republic of Belarus.

***Keywords*:** COVID-19 typology; cluster analysis; econometric modeling; statistical ratings; economic indicators.

Introduction

The problem of analyzing the COVID-19 pandemic in various aspects is given considerable attention in the world scientific literature [1]. An important direction in the ongoing research is the development of methods for statistical analysis of the COVID-19 pandemic based on the data available in the mode of regular updating. Both analytic simulation [2] and statistical models [3] are used to analysis and predict the epidemic process at the level of individual countries. Considerable attention is paid to the tasks of analyzing the COVID-19 pandemic in a multi-country aspect [1, 4].

This research has the following objects: 1) statistical multi-country analysis of the pandemic COVID-19 typology and evaluation of its influence on the economic indicators of countries by means the machine learning algorithms; 2) econometric modeling and forecasting of the epidemic process in the Republic of Belarus.

1. The problems and used data

Multy-country COVID analysis problem. It is assumed that the available panel statistical data include the values of N indicators of epidemic process obtained for some sample of countries of volume n at time $t\,(t=1,...,T)$:

$$x_{i,t}=(x_{i1,t},...,x_{iN,t})'\in\Re^N\ (i=1,...,n,\ t=1,...,T).$$

In the context of the COVID-19 analysis, panel data have a heterogeneous

cluster structure. It is supposed that the most important factor of heterogeneity is the difference between countries by a latent feature, which characterizes the intensity of the COVID-19 epidemic process. According to this property, countries can be assigned to one of the L classes. This property is expressed by a discrete random variable $d_{it} \in \{1,...,L\}$, indicating the class number for country i at time t. Class numbers $\{d_{it}\}$ are interpreted as country ratings of the intensity of the epidemic process.

The problem of statistical classification: to divide the sample $\{x_{i,t}\}$, heterogeneous in terms of latent feature, into L homogeneous subsamples (classes) that differ in the space of classification features by the degree of intensity of the epidemiological process. The solution to this problem is the classification matrix $D = \{d_{i,t}\} (i = 1,...,n,\ t = 1,...,T)$.

Single-country COVID analysis problem. The purpose of statistical analysis of the epidemic process within a single country is to solve the following tasks based on available statistical data: assessment and short-term forecasting of the growth of new infections; building long-term forecasts, the purpose of which is to assess the turning point of the epidemic wave and the moment of its completion.

To solve these problems, we use the daily and weekly data for 30 countries of the European region (Armenia, Austria, Azerbaijan, Belarus, Bulgaria, Croatia, Czech, Denmark, Estonia, Finland, France, Georgia, Germany, Great Britain, Greece, Hungary, Ireland, Italy, Kazakhstan, Latvia, Lithuania, Moldova, Netherlands, Poland, Romania, Russia, Slovenia, Spain, Switzerland, Turkey) from March 1, 2020 to April 18, 2022.

The list of available indicators includes: total number of infections (Total – $T(t)$), number of active cases of infection (*A*ctive – I(t)), number of recovered (*R*ecovered – $R(t)$), number of deaths (Deceased – $D(t)$) [5].

2. Used approaches and algorithms

Statistical analysis of the pandemic COVID-19 typology. The following classification features are constructed and used:

- the ratio of the number of closed cases to the total number of infected (Closed to Total);
- the ratio of the number of closed cases to the number of active cases (Closed to Active);
- daily growth rate of the total number of infections or the ratio of the current value of cases to the previous one (Total Infections Daily Rate);

- mortality rate – the proportion of deaths from the total number of officially registered cases of COVID-19 (Death Rate).

Since there is no training sample and the number of classes L is not known, it is necessary to use panel data classification algorithms in the self-learning mode. To solve this problem, it is proposed the approach to the analysis of panel data with a cluster structure [6].

This approach includes the following steps:

- preliminary statistical analysis of sample and outlier detection;
- censoring and scale transformation of features to interval (0,1) in such a way that values close to zero correspond to a more favorable course of the epidemic process and vice versa;
- cluster analysis of initial non-classified sample in cross-sectional representation by means of hierarchical cluster analysis and L-means algorithm;
- calculation and analysis of pandemic statistics at country and multi-country levels.

Discriminatory abilities of classification features in the L-means algorithm are shown in Figure 1.

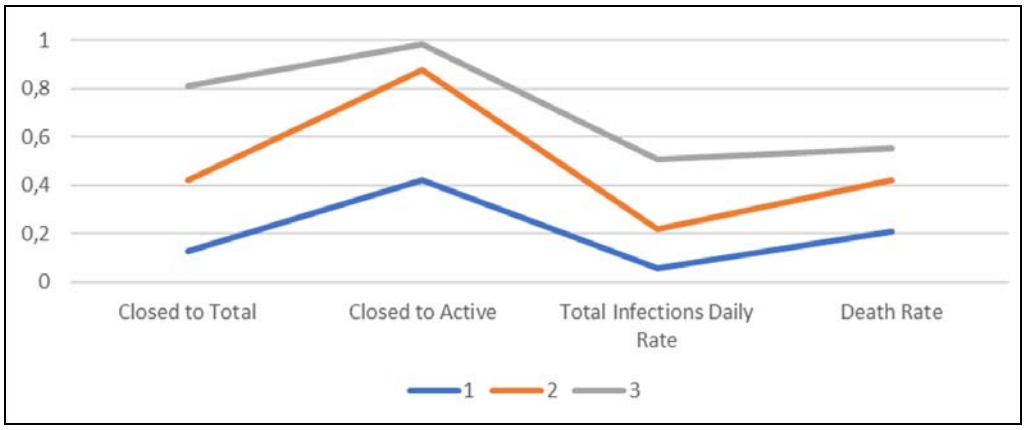


Figure 1 – Values of classification features for cluster centers 1, 2, 3

Based on the estimated classification matrix $D = \{d_{i,t}\}(i = 1,...,n,\ t = 1,...,T)$ the following indicators of the COVID-19 pandemic are constructed:

$d_{it} \in \{1, \ldots, L\}$ – *daily country rating* (*DCR*), which characterizes the degree of intensity of the epidemic for country i at time t: rating values 1 and L correspond to the lowest and highest degree of intensity of the epidemic process;

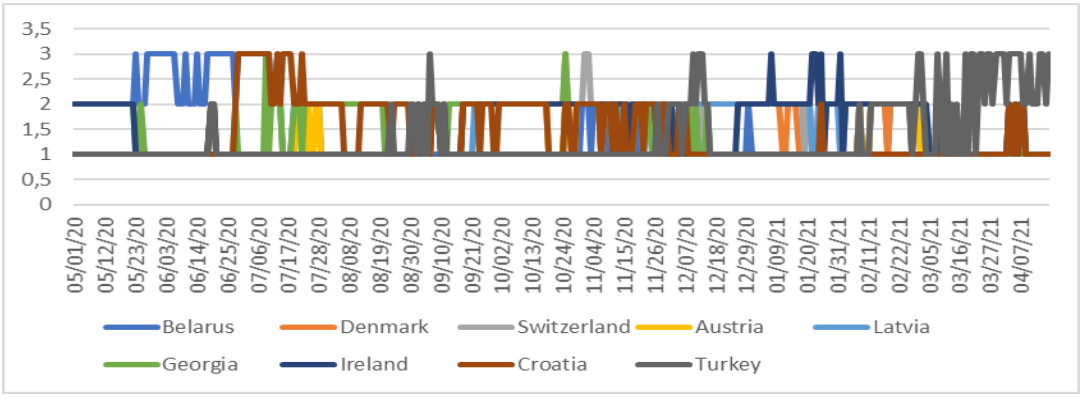
ACR_i – *Average Country Rating for the entire time interval*:

$$ACR_i = \frac{1}{T}\sum_{t=1}^{T} d_{it} \in (1,\ L),\ i = 1,...,n;$$

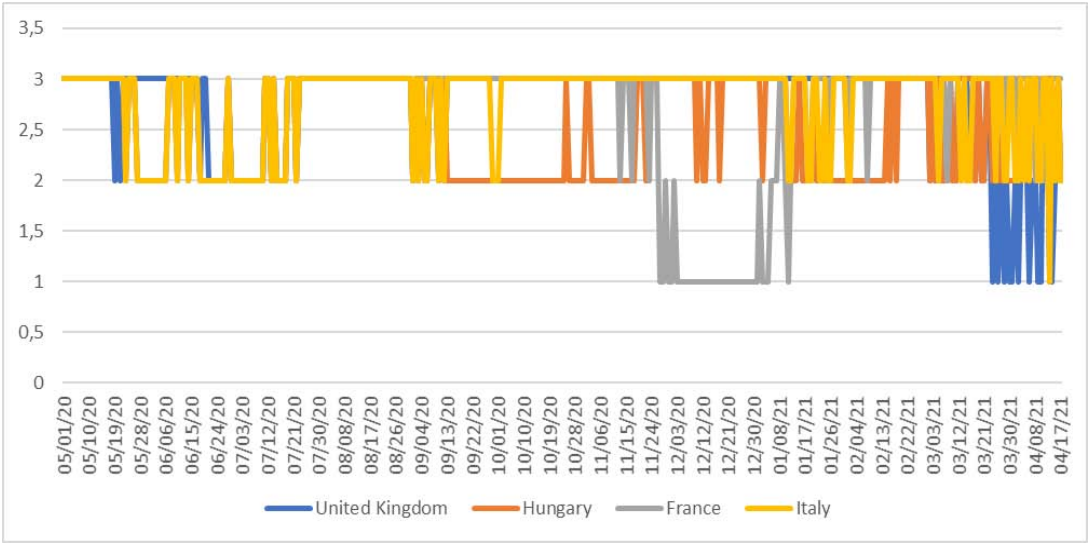
IMI_t – *Integral Multicounty Indicator of COVID-19* at time $t = 1,...,T$:

$$IMI_t = \frac{1}{n}\sum_{i=1}^{n} d_{it} \in (1,\ L),\ t = 1,...,T.$$

Figure 2 illustrates the daily DCR rating for countries from classes 1 (panel *a*) and 3 (panel *b*) with lowest and highest degree of intensity of the epidemic process respectively.



a)



b)

Figure 2 – DCR rating for countries of class 1 (a) and 3 (b) up to April 18, 2021

Table shows the average values (at the end of 2020) for GDP Annual Growth Rate and Unemployment Rate [7] for classes 1, 2, 3.

Table – Country rankings with GDP growth rates and Unemployment rate

Class (rating)	GDP Annual Growth Rate	Unemployment Rate
1	-3,610	8,395
2	-4,063	7,141
3	-7,716	6,630

The results of ranking all countries according to the ACR rating are presented on Figure 3 for two term intervals of COVID-19, including: 1) March, 2020 – April, 2021; 2) March, 2020 – April, 2022. It can be concluded that the typology of the epidemiological process in countries as a whole is preserved for new waves of the epidemic.

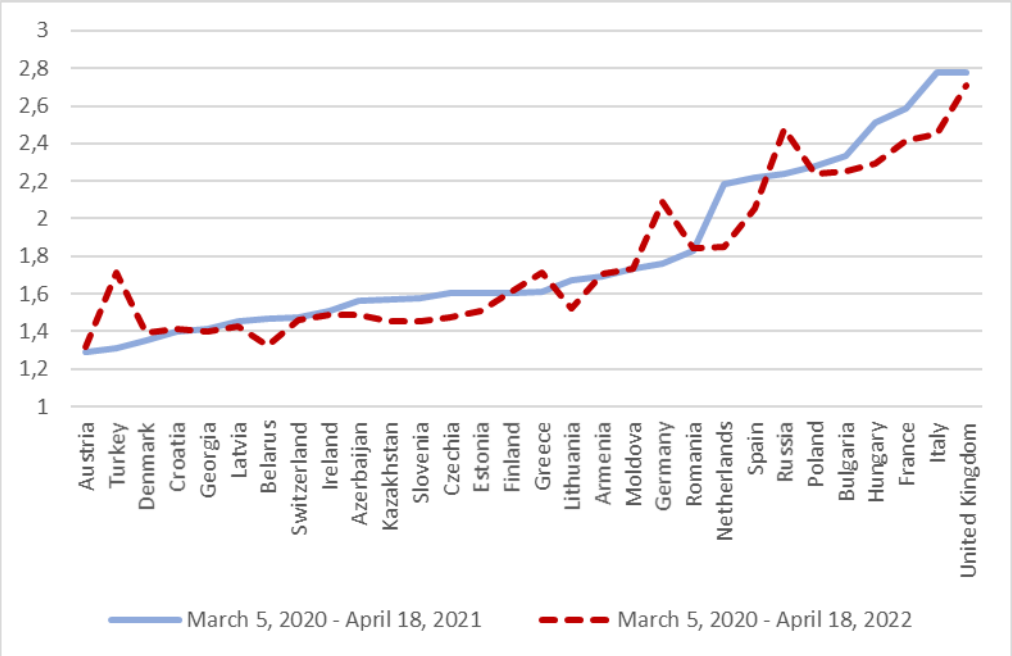


Figure 3 – The results of ranking all countries according to the ACR rating for the first two waves and entire observation period

Econometric modeling and analysis of COVID-19 in the Republic of Belarus. To analyze and predict the main indicators of the COVID-19 epidemic process two types of econometric models have been developed:

1) vector error correction model (Vector Error Correction Model – VECM COVID-19 RB) for analysis and forecasting within a single wave;

2) Markov-switching models for estimation the turning points of rise and fall of the epidemic process for the entire observation period [8].

Both models are based on assumptions close to the SIR (*Susceptible-Infectious-Recovered*) model. The main one is the assumption of the existence of a long-term equilibrium dependence for a steady state of the epidemic [2]:

$$I(t) + C(t) + S(t) = N \text{ or } I(t) + C(t) = N - S(t) = T(t),$$

where for the moment of time N – the size of the entire population; $S(t)$ – the number of persons susceptible to infection; $C(t)$ – the number of closed cases of infection, including those who recovered $R(t)$ and died $D(t)$.

Weekly time series $I(t)$, $C(t)$ are used to build a linear regression model MS-LR-AR with Markov switching of states, that allows autocorrelation of residuals. For two classes of states of the epidemic process, "rising" and "recession", the first differences of the time series $DI(t)$ and $DC(t)$ are used, that is, weekly changes in the variables $I(t)$ в зависимости от $C(t)$. The constructed model is based on the established long-term cointegration relationship between these time series and take the form:

$$DI_t = c_{d(t)} + \beta_{d(t),1} t + \beta_{d(t),2} DC_t + \eta_t,$$

where the values of the variable $d(t)$ indicate the class of epidemic states: $d(t)=1$ for the class "rising" and $d(t)=2$ for the class "decline".

All parameters of the models are unknown and are estimated using the EM (*Expectation-Maximization*) machine learning algorithm [8]. To correct the autocorrelation of residuals of the model, the algorithm proposed [9] is used.

Conclusions

Based on the obtained results of typology analysis (Figures 1–3 and Table), it can be concluded that in countries with the highest degree of intensity of the epidemic process, there is a greater decline in economic growth. It may be also supposed that the intensity of the epidemic process in each country is largely due to the ongoing anti-COVID state policy and effectiveness of anti-COVID measures.

The constructed econometric models are recommended to be used for modeling and forecasting the number of active infections within a single wave (VECM COVID-19 RB model) and for the entire period of observation of the epidemic (MS-LR-AR COVID-19 RB). The necessary condition for building these models is the stability and controllability of the epidemic process.

References

1. Cao Longbing, Liu Qing. COVID-19 Modeling: A Review // SSRN papers, 2021. URL: https://ssrn.com/abstract=3899127.

2. Kermack W., McKendrick A. Contributions to the mathematical theory of epidemics // Bulletin of Mathematical Biology. 1991. №53(1–2), pp. 33–55.
3. Kharin Yu.S, Valoshka V.A, Dernakova O.V, Malugin V.I, Kharin A.Yu. Statistical forecasting of the dynamics of epidemiological indicators for COVID-19 incidence in the Republic of Belarus // Journal of the Belarusian State University. Mathematics and Informatics. 2020. №3, pp. 36–50 (in Russian).
4. Jang S.Y., Hussain-Alkhateeb L., Rivera Ramirez, T. et al. Factors shaping the COVID-19 epidemic curve: a multi-country analysis // BMC Infect Dis 21, 2020. URL: https://doi.org/10.1186/s12879-021-06714-3
5. Worldometers.info [Electronic resource]. URL: https://www.worldometers.info/coronavirus. Date of access: 27.02.2022.
6. Malugin V.I., Hryn N.V., Novopoltsev A.Yu. Statistical analysis and econometric modelling of the creditworthiness of non-financial companies // Int. J. Computational Economics and Econometrics. 2014. Vol. 4(1/2), pp. 130-147.
7. The World Bank Group [Electronic resource]. URL: https://data.worldbank.org/. Date of access: 02.03.2022.
8. Malugin V. Statistical Estimation and Classification Algorithms for Regime-Switching VAR Model with Exogenous Variables / V. Malugin, A. Novopoltsev // Austrian Journal of Statistics. 2017. Vol. 46, pp. 47–56.
9. Malugin V.I. Discriminant analysis of multivariate autocorrelated regression observations under conditions of parametric heterogeneity of models // Informatics. 2008. №3(19), pp. 17–28 (in Russian).

УДК 519.23

МОДЕЛИРОВАНИЕ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С КОНКУРЕНЦИЕЙ ЗА ПРИБОРЫ

А.Н. Дудин, С.А. Дудин, О.С. Дудина

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, dudin@bsu.by, dudins@bsu.by, dudina@bsu.by

Рассматривается модель массового обслуживания, состоящая из двух систем, конкурирующих за приборы. Запросы двух типов поступают в соответствии с маркированным марковским входным потоком. Построен процесс изменения состояний системы. Выписан инфинитезимальный генератор данного процесса

Ключевые слова: Конкурирующие системы; маркированный марковский входной поток; многомерные цепи Маркова.

MODELING QUEUING SYSTEMS WITH COMPETITION FOR SERVERS

A.N. Dudin, S.A. Dudin, O.S. Dudina

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus, dudins@bsu.by, dudina@bsu.by
Corresponding author: dudin@bsu.by

We consider a queuing model consisting of two systems competing for servers. Two types of customers arrive according to the marked Markovian arrival process. The process of system states is analyzed. The infinitesimal generator of this process is written out.

Keywords: Competing systems; marked Markovian arrival process; multidimensional Markov chains.

Введение

Теория массового обслуживания – общепризнанный математический аппарат для оптимизации распределения ограниченных ресурсов в различных системах и сетях. Зачастую, приборы, которые не задействованы в данный момент, не простаивают, а могут быть использованы для выполнения других задач с целью получения

дополнительной прибыли. Обычно при исследовании систем с переменным числом приборов в случае необходимости подключения прибора считается, что прибор из неактивного состояния просто переключается в активное состояние. В реальности незадействованный прибор, который считается выключенным, может выполнять другую задачу, и включение его как дополнительного прибора может означать то, что прибор должен будет прекратить выполнять эту другую задачу. То есть, подключение прибора для работы в данной системе означает его выключение в другой параллельно работающей системе. В результате этого оптимальная стратегия подключения приборов может сильно отличаться от стратегии, полученной для классической системы с включением приборов.

В данной работе мы исследуем модель массового обслуживания состоящую из двух систем массвого обслуживания, конкурирующих за приборы. Каждая из систем имеет свое минимальное число приборов, которое не может быть уменьшено. Остальные приборы могут работать как в первой, так и во второй системе. Первая система считается приоритетной и может забирать приборы из второй системы.

Литература, посвященная системам обслуживания с конкуренцией за приборы довольно бедна. Похожая модель была исследована в работе [1]. В отличие от нее, где предполагалось, что запросы двух типов поступают независимо друг от друга, в данной работе мы предполагаем, что поступления запросов в обе системы зависимы и описываются маркированым марковским входным потоком (MMAP). Кроме того, мы предполагаем, что запросы находящиеся в буфере, могут проявлять нетерпеливость и покидать систему без обслуживания. Данные обобщения существенно повышают адекватность модели, но и усложняют ее аналитическое исследование.

1. Математическая модель

Мы рассматриваем модель, состоящую из двух взаимодействующих многолинейных систем массового обслуживания, конкурирующих за приборы. Структура модели представлена на рисунке.

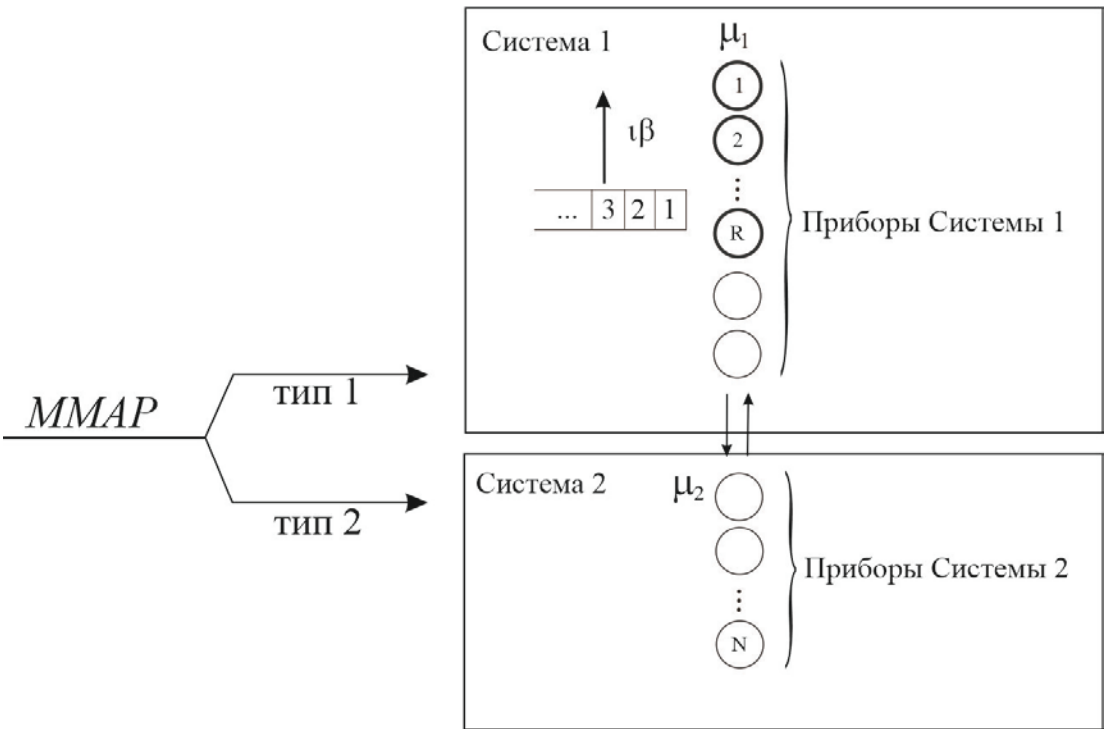


Рисунок – Структура системы

Процесс поступления запросов двух типов описывается MMAP-потоком. Данный поток задается управляющим процессом $\nu_t, t \geq 0,$ который является неприводимой цепью Маркова с непрерывным временем и конечным пространством состояний $\{1,\ldots,W\},$ и матрицами D_0, $D_1^{(1)}$ и $D_1^{(2)}$. Обозначим среднюю интенсивность поступления запросов k-го типа как λ_k. Считаем, что запросы k-го типа поступают в k-ю систему, k=1,2.

Система 1 имеет бесконечный буфер, а Система 2 не имеет буфера. Общее количество приборов в обеих системах равно N. Системы совместно используют существующие приборы следующим образом. Количество приборов, зарезервированных исключительно для использования Системой 1 (для обслуживания запросов первого типа), равно R. Количество приборов, зарезервированных исключительно для использования Системой 2, равно $M, 1 \leq M \leq N - R - 1$. Оставшийся пул из $N - R - M$ приборов может использоваться обеими системами, когда все их зарезервированные приборы заняты. Определенный приоритет в доступе к общему пулу имеет Система 1. А именно, если количество запросов, требующих обслуживания в Системе 1, невелико, то все приборы $N - R,$ которые не зарезервированы исключительно для использования Системой 1, доступны для использования Системой 2. Однако при

увеличении количества запросов в Системе 1 эта система может последовательно (по одному) забирать приборы из Системы 2. Это может привести к прекращению обслуживания запросов второго типа, получающих обслуживание на изымаемом приборе.

Правило вывода приборов Системой 1 из общего пула определяется набором порогов $(J_1, J_2, \ldots, J_{N-R-M})$, где $R < J_1 < J_2 < \ldots < J_{N-R-M}$. Если число запросов в Системе 1 меньше порога J_1, то только R приборов находятся (обслуживают или простаивают) в Системе 1. Если число заявок в Системе 1 принадлежит интервалу $[J_k, J_{k+1}), k = \overline{1, N-R-M-1}$, то $R+k$ приборов предоставляют обслуживание в Системе 1. Если число запросов в Системе 1 превышает порог J_{N-R-M}, то количество приборов, работающих в Системе 1, равно $N-M$. Если в момент, когда необходимо забрать прибор из Системы 2, в ней есть свободные приборы, то один из свободных приборов начнет работу в Системе 1. Если все приборы Системы 2 заняты и общее количество приборов позволяет изъятие прибора из нее, то один из этих приборов прекращает обслуживание и начинает обслуживание запроса типа 1 из буфера. Запрос, обслуживание которого было прекращено, теряется. Когда все приборы из общего пула обслуживают запросов первого типа, дальнейшее изъятие приборов становится невозможным. При уменьшении очереди заявок первого типа соответствующие приборы из общего пула снова становятся доступными для заявок второго типа в зависимости от соотношения текущего количества заявок первого типа в системе и пороговых значений $J_k, k = \overline{1, N-R-M}$.

Полагаем, что время обслуживания в Системе $r, r = 1, 2,$ распределено экспоненциально с параметром $\mu_r, \mu_r > 0$. Запросы первого типа, ожидающие обслуживания, могут проявлять нетерпеливость. Каждый такой запрос может покинуть буфер после экспоненциально распределенного с параметром $\beta, \beta \geq 0,$ время. В ситуации, когда из-за ухода запроса из буфера возникает необходимость передать из Системы 1 в Систему 2 прибор, которые в данный момент осуществляет обслуживание, один из занятых приборов прекращает обслуживание запроса первого типа, переходит во вторую систему, а запрос, обслуживание которого было прервано, становится на первую позицию в буфере, отодвигая стоящие там запросы.

2. Процесс изменения состояний системы

Процесс изменения состояний системы исследуемой системы можно описать регулярной неприводимой цепью Маркова с непрерывным временем

$$\xi_t = \{i_t, r_t, \nu_t\}, t \geq 0,$$

где, в момент времени t,

i_t – число запросов в Системе 1, $i_t \geq 0$;

r_t – число занятых приборов в Системе 2, $r_t = \overline{0, N-R}$, если $i_t < J_1$; $r_t = \overline{0, N-R-k}$, если $J_k \leq i_t < J_{k+1}$ и $r_t = \overline{0, M}$, если $i_t \geq J_{N-R-M}$;

ν_t – состояние управляющего процесса ММАР, $\nu_t = \overline{1, W}$.

Пронумеруем состояния цепи Маркова $\xi_t, t \geq 0$, в лексикографическом порядке. Интенсивности переходов процесса ξ_t определяются элементами его инфинитезимального генератора G.

Теорема 1. Генератор G цепи Маркова $\xi_t, t \geq 0$, имеет следующую блочно-трехдиагональную структуру

$$G = \begin{pmatrix} G_{0,0} & G_{0,1} & O & O & \dots \\ G_{1,0} & G_{1,1} & G_{1,2} & O & \dots \\ O & G_{2,1} & G_{2,2} & G_{2,3} & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Ненулевые блоки генератора определяются следующим образом:

$$G_{0,0} = I_{N-R+1} \otimes D_0 + \mu_2 C_{N-R} E_{N-R}^{-} \otimes I_W - \mu_2 C_{N-R} \otimes I_W + E_{N-R}^{+} \otimes D_2,$$

$$G_{i,i} = G_{0,0} - i\mu_1 I_{(N-R+1)W}, i = \overline{1, R},$$

$$G_{i,i} = G_{0,0} - (R\mu_1 + (i-R)\beta) I_{(N-R+1)W}, i = \overline{R+1, J_1 - 1},$$

$$G_{i,i} = I_{N-R-k+1} \otimes D_0 + \mu_2 C_{N-R-k} E_{N-R-k}^{-} \otimes I_W - \mu_2 C_{N-R-k} \otimes I_W +$$

$$+ E_{N-R-k}^{+} \otimes D_2 - ((R+k)\mu_1 + (i-(R+k))\beta) I_{(N-R-k+1)W},$$

$$i = \overline{J_k, J_{k+1} - 1}, k = \overline{1, N-R-M-1},$$

$$G_{i,i} = I_{M+1} \otimes D_0 + \mu_2 C_M E_M^- \otimes I_W - \mu_2 C_M \otimes I_W +$$

$$E_M^+ \otimes D_2 - ((N-M)\mu_1 + (i-(N-M))\beta) I_{(M+1)W}, i \geq J_{N-R-M},$$

$$G_{i,i-1} = (\min\{i,R\}\mu_1 + (i - \min\{i,R\})\beta) I_{(N-R+1)W}, 0 < i < J_1,$$

$$G_{i,i-1} = ((R+k)\mu_1 + (i-(R+k))\beta)\tilde{E}_{N-R-k}^+ \otimes I_W, i = J_k, k = \overline{1, N-R-M},$$

$$G_{i,i-1} = ((R+k)\mu_1 + (i-(R+k))\beta) I_{(N-R-k+1)W},$$
$$i = \overline{J_k + 1, J_{k+1} - 1}, k = \overline{1, N-R-M-1},$$

$$G_{i,i-1} = ((N-M)\mu_1 + (i-(N-M))\beta) I_{(M+1)W}, i > J_{N-R-M},$$

$$G_{i,i+1} = I_{N-R+1} \otimes D_1, i = \overline{0, J_1 - 2},$$

$$G_{i,i+1} = \tilde{E}_{N-R-k+1}^- \otimes D_1, i = J_k - 1, k = \overline{1, N-R-M},$$

$$G_{i,i+1} = I_{N-R-k+1} \otimes D_1, i = \overline{J_k, J_{k+1} - 2}, k = \overline{1, N-R-M-1},$$

$$G_{i,i+1} = I_{M+1} \otimes D_1, i \geq J_{N-R-M},$$

где

$\otimes$ – символ Кронекера произведения матриц, см. [2];

$C_i = \text{diag}\{0,1,\ldots,i-1,i\}$, где diag{…} обозначает диагональную матрицу, диагональные элементы которой указаны в скобках;

E_k^- – квадратная матрица размера $k+1$ со всеми нулевыми элементами, кроме элементов $(E_k^-)_{l,l-1} = 1, l = \overline{1,k}$;

E_k^+ – квадратная матрица размера $k+1$ со всеми нулевыми элементами, кроме элементов $(E_k^+)_{l,l+1} = 1, l = \overline{0,k-1}$ и $(E_k^+)_{k,k} = 1$;

$\tilde{E}_k^-$ – матрица размера $(k+1)\times k$ со всеми нулевыми элементами, кроме элементов $(\tilde{E}_k^-)_{l,l} = 1, l = \overline{0,k-1}$ и $(\tilde{E}_k^-)_{k,k-1} = 1$;

$\tilde{E}_k^+$ – матрица размера $(k+1)\times(k+2)$ со всеми нулевыми

элементами, кроме элементов $(\tilde{E}_k^+)_{l,l} = 1, l = \overline{0,k}$.

Доказательство теоремы проводится путем анализа всех вариантов переходов цепи Маркова $\xi_t, t \geq 0,$ на интервале бесконечно малой длины.

Библиографические ссылки

1. Lee, S., Dudin, A., Dudina, O., Kim, C. Analysis of a priority queueing system with the enhanced fairness of servers scheduling // Journal of Ambient Intelligence and Humanized Computing, 2022. P. 1–13.
2. Graham, A. Kronecker Products and Matrix Calculus with Applications. Cichester: Ellis Horwood. 1981. 130 p.

УДК 519.23

НАХОЖДЕНИЕ ХАРАКТЕРИСТИК ПРОИЗВОДИТЕЛЬНОСТИ МОДЕЛИ ДВУХ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ С КОНКУРЕНЦИЕЙ ЗА ПРИБОРЫ

А.Н. Дудин, С.А. Дудин, О.С. Дудина

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, dudin@bsu.by, dudins@bsu.by, dudina@bsu.by

В первой части данного исследования построена модель массового обслуживания, состоящая из двух систем, конкурирующих за приборы. Выписан инфинитезимальный генератор процесса изменения состояний системы. В данной статье для исследуемой модели найдено условие существования стационарного режима. Обсуждены способы нахождения стационарного распределения вероятностей ее состояний. Приведены формулы для нахождения основных характеристик производительности.

Ключевые слова: Конкурирующие системы; маркированный марковский входной поток; многомерные цепи Маркова; условие эргодичности.

PERFORMANCE EVALUATION OF THE MODEL CONSISTING OF QUEUING SYSTEMS WITH COMPETITION FOR SERVERS

A.N. Dudin, S.A. Dudin, O.S. Dudina

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus, dudins@bsu.by, dudina@bsu.by
Corresponding author: dudin@bsu.by

In the first part of research, we created a model consisting of two queuing systems competing for servers. The infinitesimal generator of the process of the system states has been written down. In this paper, the ergodicity condition is found. The possibilities of finding a stationary probability distribution are discussed. Formulas for calculating the main performance characteristics are given.

Keywords: Competing systems; marked Markovian arrival process; multidimensional Markov chains; ergodicity condition.

Введение

В работе [1] была построена модель массового обслуживания, состоящая из двух систем, конкурирующих за приборы. Запросы двух типов поступают в соответствии с маркированным марковским входным потоком. Процесс изменения состояний системы был задан многомерной

цепью Маркова с непрерывным временем. Выписан инфинитезимальный генератор данной цепи Маркова.

В данной статье мы продолжаем исследования, проведенные в работе [1]. Найдено условие существования стационарного режима. Обсуждены способы нахождения стационарного распределения вероятностей, а также приведены основные формулы для нахождения стационарных характеристик производительности исследуемой модели. Далее будут использованы обозначения введенные в статье [1].

1. Условие существования стационарного режима и стационарные вероятности системы

Одним из важных этапов в исследовании модели массового обслуживания является определение условия существования стационарного режима. Сперва рассмотрим случай, когда запросы первого типа являются нетерпеливыми, т.е. параметр β строго больше нуля. В данном случае, исследуемая цепь Маркова $\xi_t, t \geq 0,$ принадлежит к классу ассимптотически квазитеплицевых цепей Маркова, см. [2]. Воспользовавшись результатами из работы [2], можно формально доказать интуитивно очевидный факт, что если запросы первого типа нетерпеливы, то стационарное распределение системы существуют при любых значениях других параметров системы. Простыми словами, условие существование стационарного режима системы является условием, при котором система функционирует так, что при большом числе запросов, находящихся в ней запросы уходят из системы в среднем быстрее, чем поступают. Действительно, если запросы в буфере нетерпеливы, то там не может скопиться неограниченное число запросов. При любой положительной интенсивности нетерпеливости существует такое достаточно большое число A запросов в буфере, при котором суммарная интенсивность ухода запросов из-за нетерпеливости $A\beta$ будет превышать интенсивность входного потока запросов первого типа λ_1.

Далее, рассмотрим случай, когда запросы первого типа являются терпеливыми, т.е. $\beta = 0.$ В данном случае, при $i \geq J_{N-R-M}$ блоки генератора $G_{i,i}$, $G_{i,i-1}$ и $G_{i,i+1}$ будут иметь вид:

$$G_{i,i} = G^0 = I_{M+1} \otimes D_0 + \mu_2 C_M E_M^- \otimes I_W - \mu_2 C_M \otimes I_W +$$

$$E_M^+ \otimes D_2 - (N-M)\mu_1 I_{(M+1)W}, i \geq J_{N-R-M},$$

$$G_{i,i-1} = G^- = (N-M)\mu_1 I_{(M+1)W}, i > J_{N-R-M},$$

$$G_{i,i+1} = G^{+} = I_{M+1} \otimes D_1, i \geq J_{N-R-M}.$$

То есть, блоки генератора при $i \geq J_{N-R-M}$ не зависят от параметра i. Это в свою очередь означает, что в данном случае цепь Маркова $\xi_t, t \geq 0,$ принадлежит к классу квазитеплицевых цепей Маркова, см. [3]. Необходимое и достаточное условие существования стационарного режима квазитеплицевой цепи Маркова записывается в виде:

$$\boldsymbol{x}G^{+}\boldsymbol{e} < \boldsymbol{x}G^{-}\boldsymbol{e}, \tag{1}$$

где вектор $\boldsymbol{x}$ является единственным решение следующей системы

$$\boldsymbol{x}(G^{-} + G^{0} + G^{+})\boldsymbol{e} = \boldsymbol{0},$$

$$\boldsymbol{x}\boldsymbol{e} = 1.$$

После некоторых алгебраических преобразований можно показать, что неравенство (1) может быть преобразовано в неравенство:

$$\lambda_1 < \mu_1(N - M). \tag{2}$$

Условие эргодичности (2) также является интуитивно понятным. В случае, если запросы первого типа терпеливы для того, чтобы в системе не накапливалась бесконечно большая очередь, необходимо и достаточно, чтобы суммарная средняя интенсивность обслуживания всеми доступными для Системы 1 приборами превышала среднюю интенсивность поступления запросов первого типа.

Далее считаем, что условие эргодичности системы выполнено, то есть, существуют пределы

$$\pi(i,r,\nu) = \lim_{t\to\infty} P\{i_t = i, r_t = r, \nu_t = \nu\}.$$

Перенумеруем эти вероятности в соответствии с введенным лексикографическим порядком состояний цепи Маркова ξ_t и сформируем из них векторы-строки

$$\boldsymbol{\pi}(i,r) = (\pi(i,r,1), \pi(i,r,1), \ldots, \pi(i,r,W),$$

$$\boldsymbol{\pi}_i = \boldsymbol{\pi}(i) = (\boldsymbol{\pi}(i,0), \boldsymbol{\pi}(i,1), \ldots, \boldsymbol{\pi}(i,R_i)),$$

где

$$R_i = \begin{cases} N - R, & \text{если } i < J_1, \\ N - R - k, & \text{если } J_k \leq i < J_{k+1}, \\ M, & \text{если } i \geq J_{N-R-M}. \end{cases}$$

Стационарные вероятности цепи Маркова ξ_t находится как единственное решение системы

$$\boldsymbol{\pi} G = \mathbf{0},$$

$$\boldsymbol{\pi} \boldsymbol{e} = 1,$$

где вектор $\boldsymbol{\pi}$ определяется как

$$\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots).$$

В случае, когда запросы первого типа терпеливы, и цепь Маркова $\xi_t, t \geq 0,$ принадлежит к классу квазитеплицевых цепей Маркова, решение данной системы осуществляется стандартными методами, см., например, [3]. В противном случае, цепь Маркова $\xi_t, t \geq 0,$ принадлежит к классу ассимптотически квазитеплицевых цепей Маркова, и решение данной системы не может быть осуществлено стандартными методами. Для решения этой бесконечной системы мы рекомендуем использовать численно устойчивый алгоритм, разработанный в [4].

2. Показатели эффективности системы

Найдя стационарное распределение состояний системы, мы можем вычислить основные характеристики ее производительности.

Среднее количество запросов в обеих системах вычисляется как

$$L = \sum_{i=0}^{\infty} \sum_{r=0}^{R_i} (i + r) \boldsymbol{\pi}(i, r) \boldsymbol{e}.$$

Среднее количество запросов в Системе 1 находится как

$$L_1 = \sum_{i=1}^{\infty} i \boldsymbol{\pi}_i \boldsymbol{e}.$$

Среднее количество запросов в Системе 2 вычисляется как

$$L_2 = \sum_{i=0}^{\infty} \sum_{r=1}^{R_i} r \boldsymbol{\pi}(i, r) \boldsymbol{e}.$$

Среднее количество приборов в Системе 1 находится как

$$N_{serv-1} = \sum_{i=0}^{\infty} (N - R_i) \boldsymbol{\pi}_i \boldsymbol{e}.$$

Среднее количество приборов в Системе 2 определяется как

$$N_{serv-2} = \sum_{i=0}^{\infty} R_i \boldsymbol{\pi}_i \boldsymbol{e} = N - N_{serv-1}.$$

Среднее количество занятых приборов в Системе 1 вычисляется как

$$N_{busy-1} = \sum_{i=0}^{\infty} \min\{i, N - R_i\} \boldsymbol{\pi}_i \boldsymbol{e}.$$

Среднее количество заявок в буфере Системы 1 находится как

$$N_{buffer-1} = \sum_{i=R}^{\infty} \max\{i - (N - R_i), 0\} \boldsymbol{\pi}_i \boldsymbol{e} = L_1 - N_{busy-1}.$$

Средняя интенсивность выходного потока успешно обслуженных заявок из Системы 1 находится как

$$\lambda_{out-1} = \mu_1 N_{busy-1}.$$

Средняя интенсивность выходного потока успешно обслуженных заявок из Системы 2 вычисляется как

$$\lambda_{out-2} = \mu_2 L_2.$$

Вероятность потери заявки в Системе 2 по прибытии

$$P_{ent} = \frac{1}{\lambda_2} \sum_{i=0}^{\infty} \boldsymbol{\pi}(i, R_i) D_2 \boldsymbol{e}.$$

Вероятность потери заявки в Системе 2 из-за принудительного прекращения обслуживания определяется как

$$P_{force} = \frac{1}{\lambda_2} \sum_{k=1}^{N-R-M} \boldsymbol{\pi}(J_k - 1, N - R - k + 1) D_1 \boldsymbol{e}.$$

Вероятность потери произвольного запроса в Системе 2 вычисляется как

$$P_{loss} = P_{ent} + P_{force} = 1 - \frac{\lambda_{out-2}}{\lambda_2}.$$

Вероятность потери заявки в Системе 1 из-за нетерпеливости вычисляется как

$$P_{imp} = \frac{\beta}{\lambda_1} \sum_{i=R}^{\infty} \sum_{r=1}^{R_i} (i - (N - R_i)) \boldsymbol{\pi}(i, r) \boldsymbol{e} = 1 - \frac{\lambda_{out-1}}{\lambda_1}.$$

Заключение

Исследована модель массового обслуживания, состоящая из двух систем, конкурирующих за приборы.Такая модель адекватно описывает, например, функционирование соты сети когнитивного радио. Построен процесс изменения состояний системы. Найдено условие существования стационарного режима системы. Обсуждены способы нахождения стационарных вероятностей состояний системы. Найдены выражения для вычисления основных характеристик производительности исследуемой модели. Эти выражения могут быть использованы для формулировки и решения задач оптимального резервирования имеющихся обслуживающих устройств.

Библиографические ссылки

1. Дудин А.Н., Дудин С.А., Дудина О.С. Моделирование систем массового обслуживания с конкуренцией за приборы // Труды Международного конгресса по информатике: информационные системы и технологии (CSIST'2022). Республика Беларусь, Минск. 27 – 28 октября 2022 г. Часть 1. 250 с.
2. Klimenok V.I., Dudin A.N. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory // Queueing System, 2006. № 54. P. 245–259.
3. Neuts M. Structured Stochastic Matrices of M/G/1 Type and Their Applications. New York: Marcel Dekker. 1989.
4. Dudin S., Dudina O. Retrial multi-server queuing system with PHF service time distribution as a model of a channel with unreliable transmission of information // Applied Mathematical Modelling. 2019. T. 65. C. 676–695.

УДК 519.23

СИСТЕМА ОБСЛУЖИВАНИЯ С БЕСКОНЕЧНЫМ БУФЕРОМ И ДИСЦИПЛИНОЙ ЛИМИТИРОВАННОГО РАЗДЕЛЕНИЯ ПРОЦЕССОРА

А.Н. Дудин, С.А. Дудин, О.С. Дудина

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, dudin@bsu.by, dudins@bsu.by, dudina@bsu.by

Мы рассматриваем систему массового обслуживания с ограниченной дисциплиной лимитированного разделения процессора и бесконечным буфером. Входной поток запросов задан марковским входным процессом. Количество одновременно обслуживаемых запросов ограничено. Процесс состояний системы является многомерным марковским процессом с интенсивностями переходов, зависящими от уровня. Получен генератор этого процесса. Найдены основные характеристики производительности системы.

Ключевые слова: Марковский входной поток; лимитированное распределение процессора; бесконечный буфер.

QUEUEING SYSTEM WITH AN INFINITE BUFFER AND LIMITED PROCESSOR SHARING DISCIPLINE

A.N. Dudin, S.A. Dudin, O.S. Dudina

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus, dudins@bsu.by, dudina@bsu.by
Corresponding author: dudin@bsu.by

We consider a queueing system with limited processor sharing discipline and infinite buffer. The arrival flow is defined by Markov arrival process. The number of customers that can be serviced simultaneously is restricted . The process of the system states is defined as a level-dependent process. The generator of this process is derived. The main performance measures of the system are obtained.

Keywords: Markov arrival flow; processor sharing; infinite buffer.

Введение

Системы массового обслуживания эффективно применяются для моделирования и оптимизации различных производственных, логистических и телекоммуникационных систем и сетей. В некоторых из таких систем запросы обслуживаются по одному в порядке, заданном дисциплиной обслуживания. Однако, зачастую запросы могут обслуживаться в системе одновременно. В этом случае рассматриваются многолинейные системы. То есть пропускная способность системы делится на несколько частей, условно называемыми приборами, и каждый прибор может обслуживать один запрос. Многолинейные системы массового обслуживания являются популярным объектом для исследования. Обзор современного состояния вопроса может быть найден, например, в [1]. Стоит отметить, что многолинейные системы имеют свои недостатки с точки зрения оптимального использования ресурса системы. Например, в ситуации, когда на обслуживания находится один запрос, а приборов много, то основная часть пропускной способности не используется. Как альтернатива многолинейным системам, рассматриваются системы массового обслуживания с дисциплиной разделения процессора. Для обзора работ по системам с разделением процессора, см., например, [2, 3, 4]. Данная дисциплина предполагает, что весь ресурс системы всегда направлен на обслуживание всех имеющихся на обслуживании запросов. То есть, даже когда на обслуживании находится один запрос, ресурс системы используется полностью.

Данная работа посвящена исследованию системы массового обслуживания с дисциплиной ограниченного разделения процессора. В отличие от классических систем, данная модель имеет следующие черты, повышающие ее адекватность современным системам. Во-первых, мы предполагаем, что каждый запрос имеет требуемую скорость обслуживания, которая не может быть превышена. В действительности, если пользователю беспроводной сети связи для работы требуется определенная пропускная способность системы, то совершенно необязательно выделять ему всю пропускную способность системы. Он просто не сможет ее использовать и не будет обслуживаться быстрее. Однако, если запросов на обслуживании становится много, и пропускной способности не хватает на обслуживание всех запросов с требуемой скоростью, то допускается уменьшение средней скорости обслуживания. Во-вторых, мы предполагаем, что число запросов на обслуживании ограничено заданным управляющим параметром. Дело в том, что если не ограничивать доступ в систему, то возможно возникновение ситуации, при

которой число запросов на обслуживании окажется настолько велико, что запросы станут обслуживаться с недопустимо малой скоростью. Кроме того, в данной работе мы предполагает, что входной поток запросов задается марковским входным потоком – MAP (от англ. – Markov arrival process), что позволяет учитывать существенные флуктуации трафика, свойственные современным телекоммуникационным сетям связи. Также, для большей адекватности модели, мы считаем, что запросы, которые не были допущены на обслуживание по приходу в систему, могут ожидать обслуживание в буфере неограниченной емкости.

1. Математическая модель

Мы рассматриваем систему массового обслуживания с бесконечным буфером и дисциплиной обслуживания разделение процессора.

Структура системы представлена на рисунке.

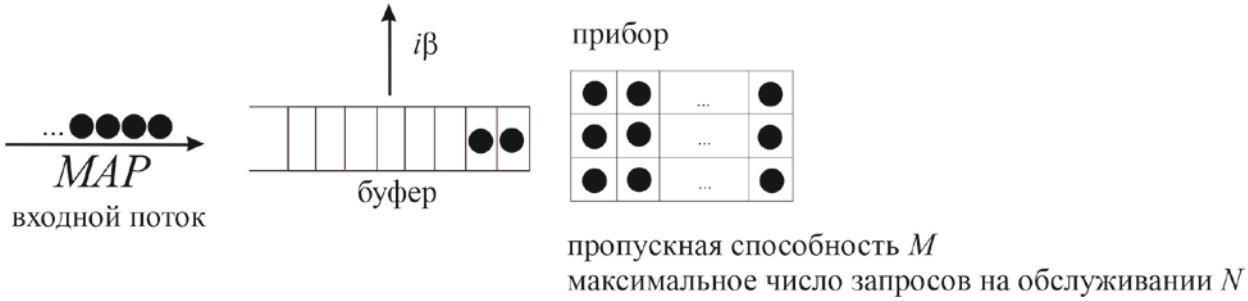


Рисунок – Структура системы

Единственный прибор может обслуживать до $N, N < \infty,$ запросов одновременно. Считаем, что общая пропускная способность прибора равна M мегабит в секунду. Для обслуживания одному запросу требуется средняя скорость X мегабит в секунду. Средний объем одного запроса составляет S мегабит. Такие образом, если запрос обслуживается с требуемой пропускной способностью, то его среднее время обслуживания определяется как $b_1 = S / X$. В данной работе будем предполагать, что время обслуживания одного запроса имеет экспоненциальное распределение. В случае наличия требуемой пропускной способности параметр экспоненциального распределения времени обслуживания задается как $\mu = 1 / b_1$. Если на обслуживании находится такое число запросов i, что $iX \le M$, что все запросы получают требуемую скорость обслуживания и обслуживаются с интенсивностью μ. В противном случае, каждому запросу выделяется пропускная способность $X_i = M / i$ мегабит и интенсивность его обслуживания равняется $\mu_i = X_i / S$.

Логично предположить, что параметр N должен быть выбран таким образом, что $NX > M$. В противном случае, пропускная способность системы не будет эффективно использоваться.

В систему поступает MAP-поток запросов. Данный поток задается управляющим процессом $\nu_t, t \geq 0,$ который является неприводимой цепью Маркова с непрерывным временем и конечным пространством состояний $\{1,\ldots,W\}$, и матрицами D_0 и D_1. Обозначим среднюю интенсивность поступления запросов как λ. Подробное описание MAP-потока, а также формулы для нахождения его характеристик можно найти в [1].

В случае, если в момент прихода запроса число запросов на обслуживании меньше параметра N, то запрос принимается на обслуживание. В противном случае, запрос идет в буфер неограниченной емкости и ожидает, пока освободится место на приборе. Запросы, ожидающие начала обслуживания в буфере, могут проявлять нетерпеливость. Это значит, что каждый запрос может покинуть буфер после экспоненциально распределенного с параметром $\beta, \beta > 0,$ времени.

2. Процесс изменения состояний системы и его анализ

Поведение рассматриваемой системы может быть описано следующей регулярной неприводимой цепью Маркова с непрерывным временем

$$\xi_t = \{i_t, \nu_t\}, t \geq 0,$$

где в момент $t, t \geq 0, i_t$ – число запросов в системе, $i_t \geq 0, \nu_t$ – состояние управляющего процесса MAP, $\nu_t = \overline{1,W}.$

Обозначим через Q генератор цепи Маркова ξ_t. Инфинитезимальный генератор Q цепи Маркова $\xi_t, t \geq 0,$ имеет блочную трехдиагональную структуру

$$Q = \begin{pmatrix} Q_{0,0} & Q_{0,1} & O & O & O & \ldots \\ Q_{1,0} & Q_{1,1} & Q_{1,2} & O & O & \ldots \\ O & Q_{2,1} & Q_{2,2} & Q_{2,3} & O & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

где ненулевые блоки $Q_{i,j}, |i-j| \leq 1,$ определяются следующим образом:

$$Q_{0,0} = D_0,\ Q_{i,i} = D_0 - i\mu I_W,\ i \le \frac{M}{X},\ Q_{i,i} = D_0 - i\mu_i I_W,\ \frac{M}{X} < i \le N,$$

$$Q_{i,i} = D_0 - (i-N)\beta I_W - N\mu_N I_W,\ Q_{i,i+1} = D_1,\ Q_{i,i-1} = i\mu I_W,\ i \le \frac{M}{X},$$

$$Q_{i,i-1} = i\mu_i I_W,\ \frac{M}{X} < i \le N,\ Q_{i,i-1} = N\mu_N I_W + (i-N)\beta I_W,\ i > N.$$

Здесь O – нулевая матрица, I – единичная матрица соответствующей размерности.

Доказательство теоремы проводится посредством тщательного анализа всевозможных переходов цепи Маркова ξ_t и последующей группировкой интенсивностей в матрицы-блоки генератора.

Исследуемая цепь Маркова ξ_t принадлежит к классу ассимптотически квазитеплицевых цепей Маркова, см. [5]. Воспользовавшись результатами из работы [5] можно формально доказать тот факт, что поскольку запросы, находящиеся в буфере проявляют нетерпеливость, то стационарное распределение системы существуют для всех значений параметров системы.

Обозначим через $\pi(i,\nu), i \ge 0, \nu_t = \overline{1,W}$, стационарные вероятности состояний цепи ξ_t. Сформируем из этих вероятностей векторы строки

$$\pi_i = (\pi(i,1),\ldots,\pi(i,W)), i \ge 0.$$

Широко известно, что вектора стационарных вероятностей π_i могут быть найдены как решение системы уравнений равновесия

$$(\pi_0, \pi_1, \ldots)\mathrm{Q} = 0, (\pi_0, \pi_1, \ldots)\mathrm{e} = 1,$$

где **e** – вектор-столбец, состоящий из единиц, и **0** – вектор-строка, состоящая из нулей.

Поскольку в рассматриваемом случае генератор Q имеет бесконечный размер, а его элементы зависят от номера строки, то решить данную систему стандартными методами не представляется возможным. Для нахождения векторов стационарных вероятностей $\pi_i, i \ge 0$, рекомендуется использовать эффективный алгоритм, разработанный в работе [6].

3. Характеристики производительности системы

Среднее число запросов на обслуживании $N_{serv} = \sum_{i=0}^{\infty} \min\{i, N\}\pi_i$e.

Среднее число запросов в буфере $N_{buffer} = \sum_{i=N+1}^{\infty} (i-N)\pi_i$e.

Среднее число запросов в системе $L = \sum_{i=0}^{\infty} i\pi_i$e $= N_{serv} + N_{buffer}$.

Вероятность того, что прибор простаивает в произвольный момент времени $P_{idle} = \pi_0$e.

Интенсивность потока обслуженных запросов

$$\lambda_{out} = \sum_{i=0}^{\infty} (\delta_{i \le \frac{M}{X}} i\mu\pi_i e + \delta_{i > \frac{M}{X}} min\{i, N\}\mu_{min\{i,N\}}\pi_i e), \text{ где } \delta_a$$

$$= \begin{cases} 1, & \text{если } a \text{ верно,} \\ 0, & \text{в противном случае.} \end{cases}$$

Вероятность того, что произвольный запрос будет потерян

$$P_{loss} = \frac{1}{\lambda} \sum_{i=N+1}^{\infty} (i-N)\beta\pi_i = 1 - \frac{\lambda_{out}}{\lambda}.$$

Вероятность того, что произвольный момент времени запросы получают урезанную скорость обслуживания

$$P_{sharing} = \sum_{i=\lceil \frac{M}{X} \rceil}^{\infty} \pi_i \text{e},$$

где $\lceil a \rceil$ определяет минимальное натуральное число большее, чем a.

Библиографические ссылки

1. Dudin A., Klimenok V. I., Vishnevsky V. M. The Theory of Queuing Systems with Correlated Flows. Cham : Springer, 2020. 430 p.
2. Yashkov S.F., Yashkova A.S. Processor sharing: A survey of the mathematical theory // Automation and Remote Control. 2007. № 68(9). P. 1662–1731.
3. Altman E., Avrachenkov K., Ayesta U. A survey on discriminatory processor sharing // Queueing systems. 2006. № 53(1). P. 53–63.

4. Kim C., Dudin S.A., Dudina O.S., Dudin A.N. Mathematical models for the operation of a cell with bandwidth sharing and moving users // IEEE Transactions on Wireless Communications. 2019. № 19(2). P. 744–755.
5. Klimenok V.I., Dudin A.N. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory // Queueing System. 2006. № 54. P. 245–259.
6. Dudin S., Dudina O. Retrial multi-server queuing system with PHF service time distribution as a model of a channel with unreliable transmission of information // Applied Mathematical Modelling. 2019. № 65. P. 676–695.

УДК 004.42+004.5

АВТОМАТИЗИРОВАННАЯ СИСТЕМА ПРИНЯТИЯ РЕШЕНИЙ ПРИ КОНТРОЛЕ ЗА ВЫБРОСАМИ ЗАГРЯЗНЯЮЩИХ ВЕЩЕСТВ

А.И. Калько, Е.А. Сундуков

Барановичский государственный универсиетет, ул. Войкова, 21, 225320, г. Барановичи, Беларусь, lexa170594@gmail.com

Рассматривается составляющая при построении автоматизированной системы принятия решений при контроле за выбросами загрязняющих веществ твердотельного котла, а именно методология и структура предметной области и ее реализация при помощи базы данных и языков программирования.

***Ключевые слова*:** Автоматизированная система; принятие решений; выброс; SCADA; датчики; база данных.

AUTOMATED DECISION-MAKING SYSTEM FOR MONITORING POLLUTANT EMISSIONS

A.I. Kalko, E.A. Sundukov

Baranavichy State University, st. Voikova, 21, 225320, Baranovichi, Belarus,
Corresponding author: lexa170594@gmail.com

The component is considered when building an automated decision-making system for monitoring pollutant emissions of a solid-state boiler, namely the methodology and structure of the subject area and its implementation using a database and programming languages.

***Keywords*:** Automated system; decision making; emission; SCADA; sensors; database.

Введение

Элементами научной новизны полученных результатов являются отсутствие физических ручных измерений, точность измерения, просмотр значений в реальном времени, автоматизированное хранение данных согласно ТКП.

Объектом исследования является автоматизированная система контроля за выбросами загрязняющих веществ в атмосферный воздух.

Предметом исследования являются методы и алгоритмы автоматизированной системы контроля за выбросами загрязняющих веществ в атмосферный воздух твёрдотопливного котла ст. №6 мини-ТЭЦ с применением сред программированиям TIA Portal и MS Virtual Studio.

Цель проекта разработать автоматизированную систему принятия решений и контроля за выбросами загрязняющих веществ в атмосферный воздух твёрдотопливного котла ст. №6 мини-ТЭЦ «Барань».

В процессе работы поставлены следующие задачи:

- изучить рабочее место оператора АСК;
- разработать методы и алгоритмы решения;
- разработать и протестировать АСК;
- разработать концептуальную, логическую и физическую модель БД.

Областью возможного практического применения являются теплоэнергостанции, котельные.

При создании АСК использованы литературные источники, научные публикации, а также самоцитирование.

1. Методология исследования / теоретические основы

Для выбора среды программирована не обходимо знать на каком ПЛК будет построена АСК, поэтому первоначально нам нужно выбрать контроллер. На данный момент всё газоаналитическое оборудование в промышленной сфере подключается/управляется с помощью дискретных или аналоговых входов/выходов на ПЛК, не редко можно встретить оборудование, которое опрашивается по «RS-485» или «RS-232» используя протокол «Modbus», но в последнее время всё чаще используется порт «RJ-45» и протоколы «Modbus RTU» и «Modbus TCP» они являются более универсальными и позволяют подключить огромное количество разных устройств с одновременным чтением и записью.

При выборе ПЛК, выбор пал на «S7-1214 DC DC RLY» бренда «SIEMENS», решение было принято из-за просто работы в среде, также в лицензии WinCC Runtime Professional на 128 тегов идёт встроенная лицензия на MS SQL Server Standart, и возможность работы с VBA-скриптами, в которых можно работать с БД [1].

Так как был выбран ПЛК «S7-1200» от бренда «SIEMENS» среда разработки будет TIA Portal V16.

TIA Portal (Totally Integrated Automation Portal) – интегрированная среда разработки программного обеспечения систем автоматизации технологических процессов от уровня приводов и контроллеров до уровня человеко-машинного интерфейса. Является воплощением концепции комплексной автоматизации (англ. Totally Integrated Automation) и эволюционным развитием семейства систем автоматизации Simatic компании Siemens AG [2].

Базовая система обладает высокой универсальностью и может использоваться в системах автоматизации различных секторов промышленного производства.

В качестве базы данных выбран MS SQL Server Standard, который идёт в пакете лицензий к WinCC Runtime Professional.

Microsoft SQL Server — система управления реляционными базами данных (РСУБД), разработанная корпорацией Microsoft. Основной используемый язык запросов — Transact-SQL, создан совместно Microsoft и Sybase [3].

Логическая модель базы данных — схема базы данных, выраженная в понятиях модели данных [4]. Логическая диаграмма автоматизированной системы состоит из далее рассматриваемых таблиц.

В таблице Value_Now хранятся текущие значения по концентрациям, выбросам, приведённые и не непреданные по «ЭкоНиП» [5], а так значение аналогового сигнала 4-20 мА по каждому компоненту (CO, CO_2, NO, NO_2, NOx и др.), а также вспомогательным датчикам (температура, расход, давление, давление воздуха КИП). Здесь же хранится слово состояние всех дискретных сигналов и HMI Alarm тегов (аварийных сообщений, уведомлений), которые для удобности сложены в три переменных типа DWORD, что равносильно типу int32 (StateWord1, StateWord2, StateWord3), благодаря преобразованию в одно слово можно записать до 32 битов (тегов типа bool) в один, который будут занимать 4 байта.

Таблицы Concentration_20m, Emission_20m, Parameter_20m хранят усреднённые значения: концентраций, выбросов и вспомогательных параметров (температура, давление, скорость и т. д.) за 20 минут соответственно, по каждой записи учитывается вид топлива, которое использовались в эти 20 минут. Архивация усреднённого значения в БД происходит строго в 19, 39 и 59 минут каждого часа, непрерывно.

Таблицы PDK_1_Day, PDK_2_DAY, PDK_3_DAY хранят предельно допустимые концентрации для каждого топлива (в основном не больше 3). Запись в каждую таблицу производится 1 раз в сутки в 0 часов 0 минут, и записанные значения будут действовать на протяжении всех суток.

Таблицы PDV_1_Day, PDV_2_DAY, PDV_3_DAY хранят предельно допустимые выбросы для каждого топлива (в основном не больше 3). Запись в каждую таблицу производится 1 раз в сутки в 0 часов 0 минут, и записанные значения будут действовать на протяжении всех суток.

Диаграмма классов автоматизированной системы состоит из далее рассматриваемых классов.

Класс ReportMounthFull отвечает за страницу месячного отчёта. Он складывает все 20 минутные значении за день и выводит таблицу средние

значения за день на заданный месяц, так же подсвечивает превышения, если они имеются.

Класс ReportDayFull отвечает за страницу суточного отчёта. Он формирует все двадцатиминутные усреднённые значения концентраций, выбросов и параметров, затем выводит их пользователю, подсвечивая значение с превышением красным цветом и процентом превышения.

Классы Value_Now, Emission_TaskVals, Concentration_20m, Emission_20m, Parameter_20m, PDK_1_Day, PDK_2_DAY, PDK_3_DAY, PDV_1_Day, PDV_2_DAY, PDV_3_DAY являются ViewModels таблиц базы данных для Entity Framework, через эти классы реализовано чтение данных из базы данных.

Классы FirstBitArray, SecondBitArray, ThirdBitArray раскладывают слово состояние из базы данных (StateWord1, StateWord2, StateWord3) на биты и присваивают значение дискретов, аварий и уведомлений переменным, для последующей удобной работы.

Класс DBContion хранит данные для подключения к базе данных автоматизированной системы.

Класс State отвечает за страницу состояние, на которых отображены все возможные состояния оборудования и уведомления, в случае аварии соответствующее состояние подсвечиваются красным цветом и включается звуковая сирена.

Класс CurrentValue отвечает за страницу значений и выводит пользователю текущие значения концентраций, выбросов, параметров приведённые и не приведённые, аналоговый сигнал 4-20 мА и предельно допустимые значения концентраций и выбросов, если они нормируются. Систематизирует соответствующей подсвечиванием компонентов и звуковой сигнализацией при превышении или подходу к пороговому значению.

Детально описываются методология и методы, на которых основано исследование.

2. Результаты и их обсуждение

Для доступа к web-приложению используется веб-браузер, желателен «Google Chrome». Для доступа на самом сервере, в адресной строке необходимо ввести http://localhost/ask или http://192.168.10.2/ask для доступа в подсети предприятия. При переходе по ссылке откроется экранная форма «Стартовая страница».

На данной странице указана общая информация о системе АСК, а также на ней расположены часы реального времени.

В верхней части экрана имеются системные состояния АСК (рисунок 1).

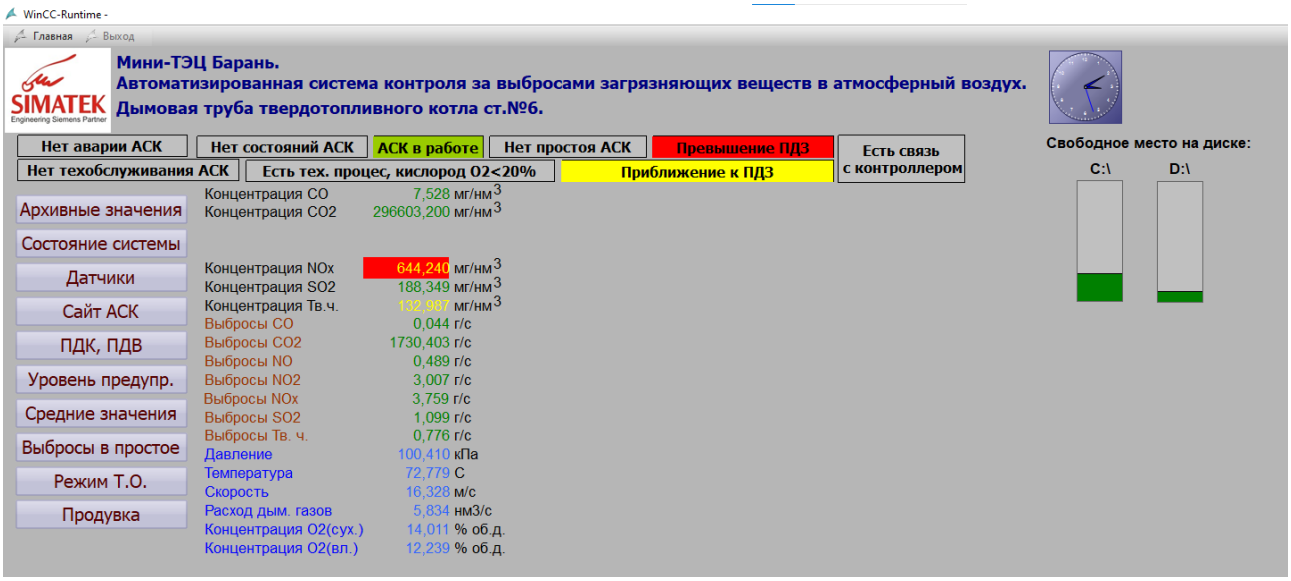


Рисунок 1 – Системные состояния АСК

Нажав кнопку «Датчики» откроется модальное окно (рисунок 2), где отображены все подключенные аналоговые датчики, текущее их значение без перерасчёта и приходящие их 4-20 мА, для аттестации аналоговых входов.

Текущие параметры датчиков

Параметр	Выходная величина		Ток (мА)
Оксид углерода (CO)	1,367	mg/m3	4,016
Диоксид углерода (CO2)	7,357	%	9,887
Оксиды азота (NOx)	162,977	mg/m3	7,477
Диоксид серы (SO2)	105,975	mg/m3	4,848
Кислород (O2), (сух. газы)	13,672	%	14,415
Кислород (O2), (влажн. газы)	11,826	%	13,106
Давление	100,498	кПа	12,346
Температура	72,562	C°	9,806
Скорость	16,092	м/с	10,436
Твердые частицы	34,447	mg/m3	6,756
Давление воздуха КИП	6,145	Bar	12,193
Температура конвертера NOx	401,628	C°	18,280
Расход пробы на газоанализаторе	80,100	литр/час	16,816

Рисунок 2 — Датчики системы

Скада WinCC Runtime после пусконаладочных работ программиста не требует запуска и выключений, она работает на протяжении всего пе-

риода работы АСК, если вдруг сервер был отключен (отключение сервера запрещено) при запуске сервера скада будет автоматически запущена.

У скады есть главное окно, которое первоначально встречает пользователя.

Данное исследование отражает предыдущие научные доклады по моделированию, проектированию, модуляции данной системы контроля за выбросами.

Заключение

Разработанный программный продукт является неотъемлемой частью сложной системы АСУ ТП предприятия и не может существовать отдельно от других, более низких по уровню компонентов системы [6]. В исследуемой работе на данном этапе отражена часть автоматизированной системы принятия решений и контроля за выбросами твердотопливного котла. Автор продолжит демонстрацию работы в последующих научных докладах.

Содержит краткие итоги разделов статьи без повторения формулировок, приведенных в них. Данный раздел может быть включен в предыдущий.

Библиографические ссылки

1. Сундуков Е.А., Калько А.И. Моделирование автоматизированной системы контроля за выбросами загрязняющих веществ // Актуальные вопросы физики и техники: материалы XI Республиканской научной конференции студентов, магистрантов и аспирантов, посвященной 100-летию со дня рождения академика Белого Владимира Алексеевича. 21 апреля 2022. Гомель, 2022. С. 237–239.
2. Интегрированная среда разработки TIA Portal v17/ SIEMENS. Германия: Siemens, 2021.
3. Работа с хранимыми процедурами SQL METANIT. URL: https://metanit.com/sql/sqlserver/11.1.php (дата обращения: 03.08.2022.)
4. Физическая и логическая модель данных. URL: https://leally.ru/good-to-know/logicheskaya-model-logicheskaya-model-dannyh-obekty-atributy-i/ (дата обращения:23.08.2022.)
5. Экологические нормы и правила Республики Беларусь. URL: https://www.ecoinfo.by/content/1327.html (дата обращения: 15.08.2022.)
6. Калько А.И., Бруйло А.А. Автоматизированное рабочее место диспетчера котельных с особыми потребностями КУП «Волковысское коммунальное хозяйство» // Непрерывное профессиональное образование лиц с особыми потребностями: сборник статей IV Международной научно-практической конференции, 9–10 декабря 2021. Минск, 2021. С. 100–103.

УДК 372+51

ОСОБЕННОСТИ ОПРЕДЕЛЕНИЯ АНОМАЛЬНЫХ ЗНАЧЕНИЙ В СИСТЕМЕ WOLRAM MATHEMATICA

Д.А. Каменко[1], М.А. Гундина[2], М.Н. Жданович[3]

[1]*Белорусский национальный технический университет,*
пр. Независимости, 65, 220013, г. Минск, Беларусь, dimakamenko.2000@gmail.com.
[2]*Белорусский национальный технический университет,*
пр. Независимости, 65, 220013, г. Минск, Беларусь, hundzina@bntu.by
[3]*«Отраслевая лаборатория новых технологий и материалов»,*
ОАО "ИНТЕГРАЛ" – управляющая компания холдинга "ИНТЕГРАЛ",
ул. Казинца, 121А, 220108, г. Минск, Беларусь, MZhdanovich@integral.by

В статье приводится обзор существующих методов определения аномальных значений выборки. Рассматривается автоматизация метода, основанного на расстоянии Махаланобиса, в компьютерной системе Wolfram Mathematica. Эта система позволяет автоматизировать процесс поиска аномальных значений выборки. Также приводится пример использования сингулярных матриц для выявления аномальных значений в выборках, представленных в виде числовых матриц.

Ключевые слова: Аномальные значения; сингулярная матрица; расстояние Махаланобиса; Wolfram Mathematica.

FEATURES OF DEFINITION OF ANOMALOUS VALUES IN WOLFRAM MATHEMATICA SYSTEM

D.A. Kamenka[a], M.A. Hundzina[b], M.N. Zhdanovich[c]

[a] *Belarusian National Technical University, 65 Nezavisimosti Ave., Minsk 220013, Belarus,* dimakamenko.2000@gmail.com.
[b] *Belarusian National Technical University, 65 Nezavisimosti Ave., Minsk 220013, Belarus,* hundzina@bntu.by
[c]*«Branch laboratory of technologies and materials», JSC "INTEGRAL" – the management company of the holding "INTEGRAL",*
121a St.Kazintsa, Minsk 220108, Belarus, MZhdanovich@integral.by

The article provides an overview of existing methods for determining the anomalous values of the sample. The automation of the method based on the Mahalanobis distance in the Wolfram Mathematica computer system is considered. This system allows you to automate the process of searching for anomalous sample values. An example of using singular matrices to detect outliers in samples presented as numerical matrices is also given.

Keywords: Anomalous values; singular matrix; Mahalanobis distance; Wolfram Mathematica.

Введение

Задача автоматизации процесса выявления аномальных значений выборки не теряет свою актуальность несколько десятилетий и находит свое применении в экономике, инженерии и других отраслях науки и техники [1–4].

Аномальные значения способны существенно исказить функционирование математических моделей анализа данных, что может привести к снижению надежности и некорректной работе всей системы, неточности прогнозов, которые будут делаться на основе таких моделей [5, 6]. Также наличие аномальных результатов может привести к недостоверным результатам при оценивании и контроле соответствия характеристик оборудования предъявляемым требованиям.

Под аномальными значениями будем понимать единицы статистической совокупности, у которых значения анализируемого признака существенно отклоняются от основного массива. Такие значения также называют выбросами.

Причины возникновения аномальных результатов разные. Такие результаты могут быть обусловлены сбоями при измерениях и регистрации данных, резкими отклонениями условий наблюдений, ошибками операторов. Поэтому необходимо выявлять и устранять аномальные результаты измерений [7].

Процесс выявления и затем удаления этих значений состоит из нескольких этапов [8]. Вначале выявляются значения, которые выходят за границы интервала возможного варьирования характеристики признака.

Исходя из физического смысла исследуемой величины, могут рассматриваться как аномальные значения те, которые не соответствуют монотонному характеру изменения величины при последовательных наблюдениях, а также значения, приращения которых превышают предельно возможную скорость изменения величины.

1. Теоретические основы

В системе *Wolfram Mathematica* есть несколько встроенных функций, которые позволяют анализировать наличие аномальных значений. Так, например, функция системы *FindAnomalies* позволяет найти члены выборки, которые считаются аномальными по отношению выборке. Задавая порог принятия можно в автоматическом режиме определять аномальные значения выборки.

Данная функция может быть использована для многих типов данных, включая числовые, строковые и графические данные.

В системе также есть и другие функции, позволяющие анализировать аномальные значения. Функция *DeleteAnomalies* выдает новый набор данных, в котором уже удалены аномальные значения. Функции *AnomalyDetection* и *AnomalyDetectorFunction* позволяют проверить, является ли новое предъявляемое значение аномальным.

Степень аномальности значения может также определяться по значению расстояния Махалонобиса. Эта величина в математической статистике является мерой расстояния между векторами случайных величин. Она обобщает понятие евклидова расстояния.

Примем в рассмотрение предположение о нормальном законе распределения исходной выборки.

Для векторизации вычисления расстояний Расстояние Махаланобиса между двумя точками – это мера расстояния между двумя случайными точками U и V, одна из которых может принадлежать некоторому классу с матрицей ковариации COV:

$$d_m(U,V,COV) = \sqrt{(U-V)COV^{-1}(U-V)^T}, \quad (1)$$

где символ *T* обозначает операцию транспонирования, а под COV^{-1} подразумевается матрица, обратная ковариационной матрице.

Элементы ковариационной матрицы вычисляются следующим образом:

$$\mathrm{cov}_{a,b} = \frac{1}{|C|-1}\sum_{x\in C}(X_1-\mu_1)(X_2-\mu_2), \quad (2)$$

где μ_1, μ_2 – математические ожидания по признакам, $|C|$ – количество точек в классе.

Расстояние Махаланобиса широко применяется в задачах кластеризации и классификации в задачах определения соответствия точки известному классу. Оно отличается от расстояния Евклида тем, что учитывает корреляции между переменными и инвариантно масштабу.

2. Результаты и их обсуждение

Метод, основанный на расстоянии Махаланобиса

Точка, имеющая наибольшее расстояние Махаланобиса до остального множества точек, считается аномалией. Такая точка имеет наибольшее влияние на кривизну и на коэффициенты уравнения регрессии. Также расстояние Махаланобиса может быть использовано в задаче определения многомерных выбросов.

Пусть исходная выборка имеет вид:

cohort={{5.04,14.22},{5.50,5.83},{5.19,4.61},{4.78,4.12},{5.08,5.99},{4.29,4.18},{5.08,6.90}};

\[Mu]=Mean@cohort;s=Covariance@cohort;

Функция ListPlot позволяет представить точки в декартовой системе координат, дополнительно указать значения расстояния Махаланобиса для каждой точки выборки:

ListPlot[{cohort,Labeled[#,Round[Dm[#,\[Mu],s],.01]]&/@points},PlotRange->All,AspectRatio->1,PlotStyle->{Darker@LightBlue,{Red,PointSize[.01]}}].

Результат работы алгоритма представлен на рис.1.

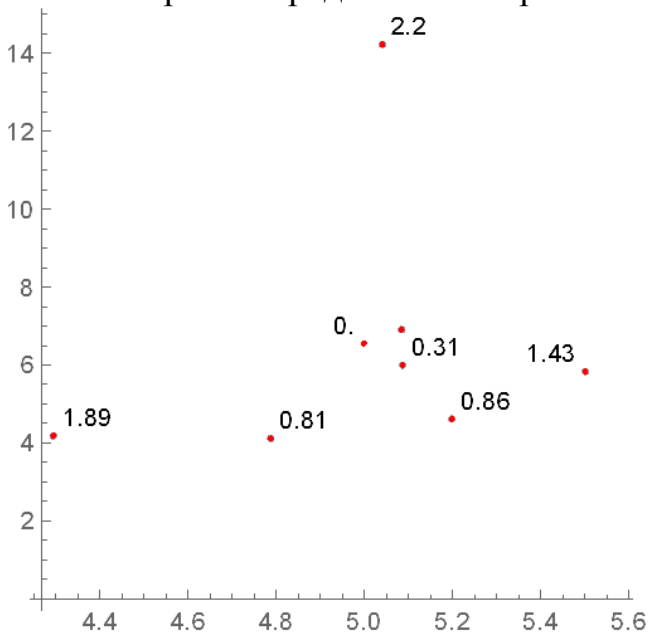


Рисунок 1 – Графическое представление исходной выборки

Анализируя полученные значения расстояний, было выявленное одно аномальное значение, величина расстояния для этого значения значительно превышает эти же расстояния для других значений выборки.

Метод, основанный на разложении матрицы

Основная идея метода обнаружения аномальных значений, основанного на разложении матриц, состоит в том, чтобы использовать сингулярное разложение исходной матрицы данных.

Наилучшая матрица получается из сингулярного разложения матрицы M по формуле:

$$M=ULV^T. \quad (3)$$

L – матрица размера mxn с неотрицательными элементами, у которой элементы, лежащие на главной диагонали – это сингулярные числа (а все элементы, не лежащие на главной диагонали, являются нулевыми), а матрицы U и V – это две унитарные матрицы, состоящие из левых и правых сингулярных векторов соответственно.

В системе Wolfram Mathematica сингулярное разложение может быть получено с помощью следующей формулы:

{u,l,v}=SingularValueDecomposition[M1];

Приближенная матрица $M_k=U_kL_kV_k^T$,U_k, L_k, V_k получаются из матриц сингулярного разложения отсечением до k первых столбцов.

Тогда приближенная матрица имеет вид, представленный на рис. 2.

$$\begin{pmatrix} 1 & 2 & 1 & 2 & 2 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 & 0 & 1 \\ 2 & 3 & 1 & 2 & 3 & 1 & 2 \\ 2 & 4 & 1 & 3 & 3 & 1 & 2 \\ 2 & 4 & 1 & 3 & 3 & 1 & 2 \\ 1 & 2 & 0 & 1 & 1 & 0 & 1 \\ 4 & 7 & 2 & 5 & 6 & 2 & 4 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 3 & 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 & 2 \\ 3 & 3 & 0 & 4 & 2 & 0 & 2 \\ 1 & 3 & 2 & 2 & 3 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 5 & 0 & 4 & 2 & 1 & 1 \end{pmatrix}$$

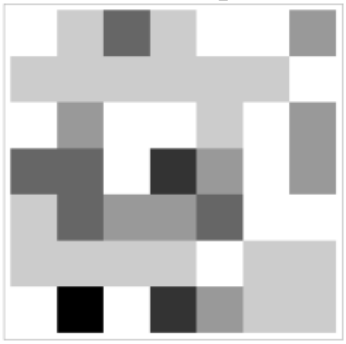


Рисунок 2 – Приближенная матрица. Матрица разности исходной матрицы с приближенной матрицей. Аномальные значения, выделенные темным цветом

Элементы, которые сильно отличаются от соответствующих элементов матрицы небольшого ранга, будут считать аномальными.

Заключение

В системе Wolfram Mathematica существуют как встроенные функции, позволяющие анализировать аномальные значения выборки, так и существует возможность автоматизировать процесс определения аномальных значений методами, основанными на расстоянии Махаланобиса и разложении матрицы.

Библиографические ссылки

1. Вероятность и математическая статистика: Энциклопедия / под ред. Ю.В. Прохорова. М.: Большая Российская энциклопедия, 2003. 911 с.
2. Линник Ю.В. Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. М.: Физматгиз, 1962. 352 с.
3. Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений. Ленинград: Энергоатомиздат, 1985. 248 с.

4. Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. М.: Наука, 1965. 552 с.
5. Богатырев В.А., Богатырев С.В. Надежность мультикластерных систем с перераспределением потоков запросов // Известия высших учебных заведений. Приборостроение. 2017. Т. 60. № 2. С. 171-177.
6. Богатырев В.А., Винокурова М.С., Петров П.А., Назарова М.Л., Шабаков Р.В. Контроль и безопасность функционирования дублированных компьютерных систем // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 2. С. 368-372.
7. Лукин В.Л., Сухорученков Б.И., Кузнецов В.И. Статистический метод выявления аномальных результатов измерений характеристик технических систем // Двойные технологии. 2010. № 2(51). С. 32-40.
8. Сухорученков Б. И., Меньшиков В. А. Методы анализа характеристик летательных аппаратов. М.: Машиностроение, 1995. 475 с.

УДК 519.21

СТАЦИОНАРНОЕ РАСПРЕДЕЛЕНИЕ СЕТЕЙ С ТРЕБОВАНИЯМИ РАЗНОГО ТИПА И ЭКСПОНЕНЦИАЛЬНЫМ ОГРАНИЧЕНИЕМ НА ВРЕМЯ ПРЕБЫВАНИЯ

В.А. Немилостивая, Ю.В. Малинковский

Гомельский государственный университет им. Ф. Скорины
ул. Кирова, 119, 246019, г. Гомель, Беларусь,
e-mail: vetta.i.am@gmail.com , Malinkovsky@gsu.by

Рассматривается экспоненциальная сеть массового обслуживания с однолинейными и многолинейными узлами. В систему поступают требования различного типа. При этом длительность ожидания обслуживания является случайной величиной, имеющей показательное распределение, зависящее от числа требований в узле. Когда требование обслужилось или время ожидания обслуживания закончилось, заявка переходит в другой узел системы и меняет свой тип или покидает сеть. Устанавливаются достаточные условия эргодичности, условия существования и вид стационарного распределения вероятностей состояний в мультипликативной форме.

Ключевые слова: сеть массового обслуживания; ограниченное время ожидания; однолинейные узлы; многолинейные узлы; стационарное распределение.

STATIONARY DISTRIBUTION FOR NETWORKS WITH DIFFERENT REQUIREMENTS AND EXPONENTIAL CONSTRAINED SOJOURN TIME

V.A. Niamilastsivaya, Yu.V. Malinkovskii

F. Scorina Gomel State University, Gomel, BELARUS,
e-mail: vetta.i.am@gmail.com , Malinkovsky@gsu.by

We consider an exponential queuing network containing nodes of two types – single-line and multi-line. The network receives customers of various types. The sojourn time of customers at the network nodes is a random value whose conditional distribution for a fixed number of customers in a node is exponential. When the requirement is serviced or the service waiting time is over, the request moves to another node of the system and changes its type or leaves the network. Sufficient conditions for ergodicity, existence conditions, and the form of the stationary distribution of state probabilities in multiplicative form are established.

Keywords: queuing network; limited waiting time; single-line nodes; multi-line nodes; stationary distribution.

Введение

Функционирование многих реальных объектов в области информационно-вычислительных систем описывают сети массового обслуживания. При этом аналитические результаты теории сетей массового обслуживания используются и при имитационном моделировании.

Сети массового обслуживания с ограничениями на время пребывания в узлах ранее рассматривались в работах [1–4]. История и суть вопроса изложены в издании Б. В. Гнеденко, И. В. Коваленко [1]. В публикации [5] исследуется стационарное поведение сети, в которой в некоторых узлах матрицы маршрутизации обслуженных заявок и заявок, время пребывания которых истекло, различны, а в других узлах – совпадают. Модель сети в данной работе усложняется тем, что требования, попадающие в узел, могут быть различных типов.

1. Постановка задачи

В сеть массового обслуживания, образованную N узлами, Q из которых однолинейны, а о остальные $N-Q$ многолинейны, поступает стационарный пуассоновский поток с интенсивностью λ. Поступающее требование независимо от других направляется в i-й узел и становится требованием типа l с вероятностью $p_{0(i,l)}$ $\left(\sum_{i=1}^{N}\sum_{l=1}^{M} p_{0(i,l)} = 1\right)$. Система с бесконечной очередью ожидания. Многолинейные системы можно свести к однолинейным с переменной условной интенсивностью обслуживания $\mu(n) = \mu \mathrm{I}_{\{n\neq 0\}}$, где n – количество требований в системе (I_A – индикатор события A, равный 1, если A происходит, и равный 0, если A не происходит).

Время обслуживания требования в i- ом однолинейном узле имеет показательное распределение с параметром μ_i $\left(i = \overline{1,Q}\right)$, а условное распределение времени обслуживания требования в остальных $N-Q$ узлах, при наличии в узле n_i требований – показательное с параметром $\mu_i(n_i)$, при этом $\mu_i(n_i) > 0$ для $n_i \in \mathrm{N}$ и $\mu_i(0) = 0$ $\left(i = \overline{Q+1, N}\right)$.

Длительность пребывания требования в i- ом узле – случайная величина, имеющая показательное условное распределение с параметром $\frac{\nu_i}{n_i}$ $\left(i = \overline{1, N}\right)$ при условии что в i-м узле находится n_i заявок. Если требование пребывает в свободный узел, оно сразу начинает обслуживаться. Требования обслуживаются в порядке поступления в узлы.

Требование типа l, завершившее обслуживание в i-м узле, моментально и независимо от других требований, переходит в j-й узел сети и становится требованием типа m с вероятностью $p_{(i,l)(j,m)}$, а с вероятностью $p_{(i,l)0}$ покидает сеть $\left(i, j=\overline{1,N},\ l,m=\overline{1,M},\ \sum_{j=1}^{N}\sum_{m=1}^{M} p_{(i,l)(j,m)} + p_{(i,l)0} = 1\right)$. Требование типа l, время пребывания которого в i-м узле завершилось (для узлов $i=\overline{1,Q}$) моментально и независимо от других требований, переходит в j-й узел и становится требованием типа m с вероятностью $r_{(i,l)(j,m)}$, а с вероятностью $r_{(i,l)0}$ покидает сеть $\left(i=\overline{1,Q}, j=\overline{1,N},\ l,m=\overline{1,M},\ \sum_{j=1}^{N}\sum_{m=1}^{M} r_{(i,l)(j,m)} + r_{(i,l)0} = 1\right)$. Если же $i=\overline{Q+1,N}$, то требование поступает как завершившее обслуживание, то есть с вероятностью $p_{(i,l)(j,m)}$ направляется в j-й узел и становится требованием типа m, а с вероятностью $p_{(i,l)0}$ покидает сеть $\left(j=\overline{1,N}, l,m=\overline{1,M}\right)$. Для удобства введём ещё узел 0, отождествляющий внешность сети.

Стохастические матрицы $P=\left\|p_{(i,l)(j,m)}\right\|$ $\left(i,j=\overline{0,N},\ l,m=\overline{0,M}\right)$ и $R=\left\|r_{(i,l)(j,m)}\right\|$ $\left(i=\overline{0,Q}, j=\overline{0,N},\ l,m=\overline{0,M}\right)$, где $p_{(0,0)(0,0)}=r_{(0,0)(0,0)}=0$, $p_{(0,0)(i,l)}=r_{(0,0)(i,l)}=r_{0(i,l)}=p_{0(i,l)}$, можно рассматривать как неприводимые марковские цепи, состояния которых обозначаются парами (i,l). Матрица P является как матрицей маршрутизации обслуженных заявок, так и матрицей маршрутизации неудовлетворенных заявок для узлов $i=\overline{Q+1,N}$, а $R-$ матрицей маршрутизации неудовлетворенных заявок для узлов $i=\overline{1,Q}$. Стахостическая матрица маршрутизации, которая управляет движением требований по узлам $i=\overline{0,N}$, без учёта того, за счёт чего требование покидает сеть (обслуживание или окончание длительности пребывания) $S=\left\|s_{(i,l)(j,m)}\right\|, \left(i,j=\overline{0,N}, l,m=\overline{0,M}\right)$, где для $(i,l)\neq 0$

$$s_{0(j,m)}=p_{0(j,m)},\ s_{(i,l)(j,m)}=\frac{\mu_i p_{(i,l)(j,m)}+\nu_i r_{(i,l)(j,m)}}{\mu_i+\nu_i}=\frac{\mu_i}{\mu_i+\nu_i}p_{(i,l)(j,m)}+\frac{\nu_i}{\mu_i+\nu_i}r_{(i,l)(j,m)},$$

для $i=\overline{1,Q}$, и $s_{(i,l)(j,m)}=p_{(i,l)(j,m)}$ для $i=\overline{Q+1,N}$.

Состояние сети описывается вектором $x=\left(x_1,x_2,\ldots,x_N\right)$, где $x_i=\left(x_{i1},x_{i2},\ldots,x_{in_i}\right)$ – вектор переменной размерности, а $x_{ij}\left(1\le j\le n_i, 1\le x_{ij}\le M, n_i=0,1,\ldots\right)$ – это тип требования, которое занимает j-ое место в очереди i- ого узла. Первым обслуживается требование x_{i1},

остальные ожидают в очереди. Последним на прибор попадет требование x_{in_i}. Причем $x_i = 0$ если $n_i = 0$, т.е. система пустая.

Будем использовать обозначения $p(x)$– вероятность того, что в момент времени t состояние сети $x = (x_1, x_2, \ldots, x_N)$; $[\tilde{x}_i]$ – вектор, все компоненты которого совпадают с вектором $x = (x_1, x_2, \ldots, x_N)$, а i-ая компонента равна $\tilde{x}_i$; $[\tilde{x}_i, \tilde{x}_j]$– вектор, все компоненты которого совпадают с вектором $x = (x_1, x_2, \ldots, x_N)$, а i-ая и j-ая компоненты равны $\tilde{x}_i$ и $\tilde{x}_j$. Введём также оператор $\mathrm{T} : (x_{i1}, x_{i2}, \ldots, x_{in_i}) \to (x_{i1}, x_{i2}, \ldots, x_{in_i - 1})$ и оператор $\mathrm{K}_l : (x_{i1}, x_{i2}, \ldots, x_{in_i}) \to (l, x_{i1}, x_{i2}, \ldots, x_{in_i})$.

2. Изолированный узел

Рассмотрим изолированно от сети i-ый узел, полагая, что в него поступает M независимых пуассоновских потоков требований интенсивности $\lambda\varepsilon_{i,l}$, $l = \overline{1, M}$. В остальном, касающемся процессов обслуживания и ограничениях на длительность пребывания, поведение изолированного узла такое же, как и в сети.

Для эргодичности цепи Маркова $x_i = (x_{i1}, x_{i2}, \ldots x_{in_i})$, описывающей изолированный узел, необходимо чтобы загрузка i-ого узла $\rho_i = \dfrac{\lambda}{\mu_i + \nu_i} \varepsilon_{i, x_{il}} < 1, \quad i = \overline{1, Q}.$

Проверено, что

$$p_i(x_{i1}, x_{i2}, \ldots x_{in_i}) = \frac{\lambda^{n_i}}{(\mu_i + \nu_i)^{n_i}} \prod_{\kappa=1}^{n_i} \varepsilon_{i x_{i\kappa}} p_i(0), \quad n_i = 1,2,\ldots,$$
$$p_i(0) = \left[\sum_{n_i=0}^{\infty} \prod_{\kappa=1}^{n_i} \left(\varepsilon_{i x_{i\kappa}} \frac{\lambda}{\mu_i + \upsilon_i} \right) \right]^{-1}. \qquad (1)$$

является стационарным распределением цепи Маркова $x_i = (x_{i1}, x_{i2}, \ldots x_{in_i})$.

Теперь изолируем i-ый узел для $i = \overline{Q+1, N}$. Имеем систему с поступающим пуассоновским потоком интенсивности $\lambda \sum_{l=1}^{M} \varepsilon_{i,l}$, а условное распределение длительности обслуживания прибором при условии, что в системе находится n_i заявок, – показательное с параметром $\mu_i(n_i)$, зависящим от n_i. Время пребывания требования в узле – случайная величина,

условное распределение которой при фиксированном n_i – показательное с параметром $\frac{\nu_i}{n_i}$.

Стационарное распределение цепи Маркова $x_i = (x_{i1}, x_{i2}, \dots x_{in_i})$, описывающей изолированный узел

$$p_i(x_{i1}, x_{i2}, \dots x_{in_i}) = \lambda^{n_i} \prod_{\kappa=1}^{n_i} \left(\frac{\varepsilon_{ix_{i\kappa}}}{\mu_i(k) + \nu_i} \right) p_i(0), \quad n_i = 1,2,\dots,$$
$$p_i(0) = \left[\sum_{n_i=0}^{\infty} \prod_{\kappa=1}^{n_i} \left(\varepsilon_{ix_{i\kappa}} \frac{\lambda}{\mu_i(k) + \upsilon_i} \right) \right]^{-1}. \tag{2}$$

Для эргодичности цепи Маркова $x_i = (x_{i1}, x_{i2}, \dots x_{in_i})$, достаточно сходимости ряда

$$\sum_{n_i=0}^{\infty} \prod_{\kappa=1}^{n_i} \left(\varepsilon_{ix_{i\kappa}} \frac{\lambda}{\mu_i(k) + \upsilon_i} \right) < \infty. \tag{3}$$

3. Основной результат

Поскольку требования при прохождении узлов не рождаются и не теряются, то в стационарном режиме выполняется следующий закон сохранения (уравнение трафика):

$$\varepsilon_{i,l} = p_{0(i,l)} + \sum_{j=1}^{N} \sum_{m=1}^{M} \varepsilon_{j,m} s_{(j,m)(i,l)}, \quad i = \overline{1,N}, \quad l = \overline{1,M}. \tag{4}$$

Здесь $\varepsilon_{i,l}$ – средняя интенсивность поступления требований типа l в i-ый узел когда сеть находится в стационарном режиме.

Доказана следующая теорема:

Теорема 1. При выполнении условия

$$\begin{cases} \dfrac{\lambda}{\mu_i + \nu_i} \varepsilon_{i,x_{il}} < 1, i = \overline{1,Q}, l = \overline{1,M}, \\ \displaystyle\sum_{n_i=0}^{\infty} \prod_{k=1}^{n_i} \frac{\lambda \varepsilon_{ix_{ik}}}{\mu_i(k) + \nu_i} < +\infty, i = \overline{Q+1,N} \end{cases} \tag{5}$$

цепь Маркова $x(t)$ эргодична, а её единственное стационарное распределение имеет форму произведения $p(x) = p_1(x_1) p_2(x_2) \dots p_N(x_N)$, где $p_i(x_i)$ –

стационарное распределение изолированного i-го узла, а $\{\varepsilon_{i,l}, i=\overline{1,N}, l=\overline{1,M}\}$ *– решение уравнения трафика (4).*

Найти различные показатели эффективности функционирования сети в стационарном режиме не составляет труда зная стационарное распределение системы.

Полученные результаты могут быть применены при проектировании новых и модернизации уже существующих сетей передачи данных и информационно-вычислительных сетей.

Библиографические ссылки

1. Гнеденко Б.В. Введение в теорию массового обслуживания. М.: Наука, 1987. 431с.
2. Ковалёв Е.А. Сети массового обслуживания с ограниченным временем ожидания в очередях // АВТ. 1985. № 2. С. 50 – 55.
3. Якубович О.В., Евдокимович В.Е. Сеть массового обслуживания со случайным временем пребывания положительных, отрицательных заявок и сигналов // Проблемы физики, математики и техники. 2010. № 4(5). С. 63 – 67.
4. Якубович О.В., Дудовская Ю.Е. Многорежимная сеть массового обслуживания со случайным временем пребывания различных типов отрицательных заявок // Проблемы физики, математики и техники. 2012. № 4(137). С. 74 – 77.
5. Малинковский Ю.В., Немилостивая В.А. Стационарное распределение сетей Джексона с экспоненциальным ограничением на время пребывания заявок // Проблемы физики, математики и техники. 2020. № 3(44). С. 73 – 77.

АНАЛИЗ СЕТЕВОГО ТРАФИКА С ПОМОЩЬЮ УСТОЙЧИВОГО РАСПРЕДЕЛЕНИЯ

Т.В. Соболева

Белорусский государственный университет
Минск, Республика Беларусь
Soboleva@bsu.by

Рассматриваются вопросы использования устойчивых распределений при анализе сетевого трафика.

Ключевые слова: самоподобие; устойчивое распределение; показатель Херста; сетевой трафик.

NETWORK TRAFFIC ANALYSIS USING SUSTAINABLE DISTRIBUTION

T.V. Soboleva

Belarusian State University
Minsk, Republic of Belarus
E-mail: Soboleva@bsu.by

The issues of using stable distributions in the analysis of network traffic are considered.

Keywords: self-similarity; sustainable distribution; Hurst exponent; network traffic.

Введение

Как известно, при проектировании сетей используют каналы связи с пакетной коммутацией, что увеличивает эффективность использования каналов, однако снижает надёжность доставки информации. Постоянное увеличения объёма передаваемых данных привело к тому, что задача прогнозирования сетевого трафика стала достаточно актуальной. Кроме того, пульсирующий характер и структурные изменения в поведении сетевого трафика могут свидетельствовать о различных атаках на сеть. Таким образом, своевременное обнаружение аномального трафика, позволяет вовремя блокировать возможную атаку.

Последние исследования показывают, что сетевой трафик, имеющий пульсирующий характер, хорошо описывается с помощью устойчивых процессов. Самоподобная структура сетевого трафика оказывает сильное влияние на производительность сети и является характерной особенно-

стью современных телекоммуникационных сетей, что делает учет описанных свойств актуальной задачей.

В статье приведены результаты статистических исследований реального сетевого трафика с помощью устойчивых распределений.

1. Основная часть

Сетевым трафиком (англ. traffic - «движение») называется объем информации, передаваемой через компьютерную сеть за определенный промежуток времени. Количество трафика может быть измерено в пакетах или таких единицах измерения как биты, байты, и их производных. Трафик подразделяется на внешний и внутренний, исходящий и входящий. [1] Мы будем использовать понятие нормального и аномального трафика. Под аномальным трафиком будем подразумевать трафик, не характерный для обычной работы сети. Такой трафик может информировать об осуществлении сетевой атаки.

О самопомоподобной структуре сетевого трафика начали говорить в начале 90-х годов, когда резко увеличился объем передаваемых данных. Ряд статей указывают на то, что объединенный из нескольких источников трафик становится сильно автокоррелированным с долговременной зависимостью [2, 3]. Такое поведение может быть объяснено тем, что будущее процесса определяется его прошлым, причем с убывающей степенью влияния этого прошлого на процесс. Совокупность множества источников данных, проявляющих синдром бесконечной дисперсии, в результате дает самоподобный объединенный сетевой трафик.

Коэффициент Херста [4, 5] является важнейшим параметром, характеризующим степень самоподобия. Этот параметр был назван в честь Х.Е. Херста –британского гидролога, который посвятил себя изучению реки Нил, а также проблеме хранения воды. Оценка параметра помогает не только решить, является ли процесс самоподобным, но и позволяет применить к процессу ряд метод по прогнозированию фрактальных процессов. Коэффициент Херста принимает значения $0 < H < 1$.

• При значении коэффициента $0.5 < H < 1$ говорят персистентном (поддерживающемся) поведении процесса, либо о том, что процесс обладает длительной памятью, то есть является самоподобным. Персистентные стохастические процессы обнаруживают четко выраженные тенденции изменения при относительно малом “шуме”.

• В случае $H = 0.5$ говорят о полностью случайном ряде, аналогичном смещениям частицы при классическом броуновском движении.

• В случае $0 < H < 0.5$ говорят о антиперсистентности процесса. Такой ряд не обладает самоподобием.

Отметим, что существуют различные методы для оценки параметра *H* для того, чтобы выявить самоподобие или медленно убывающую зависимость.

Временные методы оценки параметра Херста:

• Метод R/S статистики (В основе R/S анализа лежит формула Альберта

Эйнштейна о броуновском движении частиц).

• Метод вариаций.

• Метод абсолютного момента.

• Метод отношения вариации остатков.

• Частотные методы оценки параметра Херста.

В работе использовался метод R/S статистики.

Многочисленные исследования сетевого трафика показали, что он лучше всего описывается так называемыми распределениями с “тяжелыми хвостами”. Рассмотрим понятие устойчивого распределения. Устойчивое распределение — это распределение, которое может быть получено как предел по распределению сумм независимых случайных величин [6].

Пусть $\xi, \xi_1, \xi_2, \ldots$ — независимые, одинаково распределённые случайные величины.

Определение 1. Случайная величина ξ называется устойчивой (устойчивой в широком смысле), если для каждого m, $m = 1, 2, \ldots$, существуют такие постоянные $c_m = m^{1/\alpha}$, $\alpha \in (0,2]$, и $\gamma_m \in R$, что

$$\sum_{i=1}^{m} \xi_i \overset{d}{=} c_m \xi + \gamma_m,$$

где «$\overset{d}{=}$» означает равенство по распределению, $R = (-\infty, +\infty)$.

При исследовании устойчивых распределений удобно использовать аппарат характеристических функций. Для того, чтобы случайная величина ξ была устойчивой, необходимо и достаточно, чтобы логарифм её характеристической функции $\varphi_\xi(t)$ имел представление

$$\ln \varphi_\xi(t) = i\mu t - \sigma^\alpha |t|^\alpha + i\sigma^\alpha t \omega(t, \alpha, \beta),$$

где $\alpha \in (0, 2]$, $\beta \in [-1, 1]$, $\sigma > 0$, $\mu \in R$, $t \in R$,

$$\omega(t, \alpha, \beta) = \begin{cases} |t|^{\alpha-1} \beta tg \dfrac{\pi}{2} \alpha, & \alpha \neq 1, \\ -\beta \dfrac{2}{\pi} \ln|t|, & \alpha = 1. \end{cases}$$

Таким образом, класс устойчивых распределений представляет собой четырёхпараметрическое семейство с параметрами: α — характеристическим показателем, β — параметром асимметрии, σ — параметром масштаба, μ — параметром положения.

Данные реального сетевого трафика при больших объемах обладают свойством самоподобия, поэтому классические модели не могут быть использованы для моделирования. Для моделирования такого трафика используется устойчивое распределение.

Для исследований сетевого трафика, использовались выборки из наблюдений за трафиком организации в течении одного дня. Файлы с данными были предобработаны и подготовлены для последующего моделирования и исследования.

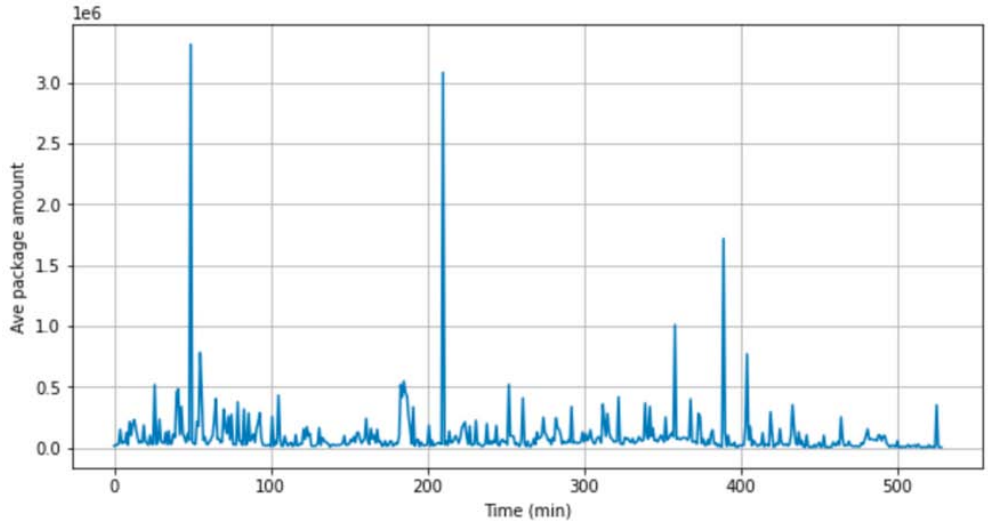
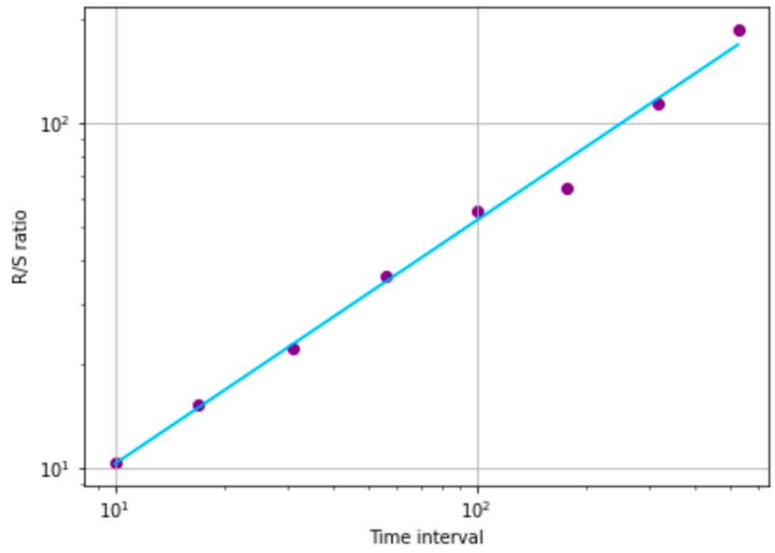


Рис.1. – График количества пакетов в зависимости от момента времени

Из рисунка 1 видно, что у нас присутствуют скачки активности, которые выходят за обычные значения.

Для оценки показателя Херста, использовался пакет «hurst» языка Python.



H=0.7046, c=2.0317

Рис.2. – Результаты R/S анализа исходного временного ряда

Как видно из рисунка 2, показатель Херста больше граничного значения 0.5, следовательно, временной ряд обладает свойством самоподобия.

Заключение

Итак, результаты исследований показали, что временной ряд, полученный в ходе преобразований исходных данных, собранных в течение одного дня обладает характеристикой самоподобия и может быть смоделирован с помощью устойчивого распределения.

Библиографические ссылки

1. Lucas M. W. Network flow analysis. N.-Y, 1990. Ch. 1. P. 9–11.
2. Jain R. Routhier S.A. Packet Trains – Measurement and a New Model for Computer Network Traffic // IEEE Journal on Selected Areas in Communications. Sep.1986. Vol. 4. № 6. P. 986–995.
3. Willinger W., Taqqu M.S., Sherman R., Wilson D.V. Self-Similarity through High Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level // IEEE/ACM Transcriptions on Networking. Feb. 1997. Vol. 5. № 1. P. 71–86.
4. Anis A.A., Lloyd E.H. The expected value of the adjusted rescaled Hurst range of independent normal summands // Biometrica. 1976. № 63. P. 283–298.
5. Simmross-Wattenberg F., Anomaly Detection in Network Traffic Based on Statistical Inference and alpha-Stable Modeling // IEEE Transactions on Dependable and Secure Computing Jul. 2001. Vol. 8. № 4. P.494–509.
6. Труш Н. Н., Соболева Т.В. Статистический анализ оценок спектральных плотностей устойчивых случайных процессов. Минск: БГУ, 2008. 100 с.

УДК 69.05:658.512.6; 355/359.07

ВЫБОР МОДЕЛИ НЕЙРОННОЙ СЕТИ В ЦЕЛЯХ СОЗДАНИЯ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ

М.Н. Сорокин, Д.С. Рябенко

Институт пограничной службы Республики Беларусь,
ул. Славинского, 4, 220103, г. Минск, Беларусь, sarokin-maksim@rambler.ru

В статье предложен вариант модели нейронной сети для разработки системы поддержки принятия решения, создаваемой для начальника подразделения пограничного контроля. Предлагаемая модель основана на системе поддержки принятия решений, ориентированной на знания. Подходы к ее реализации включают научные методы системного анализа и технологию машинного обучения.

Ключевые слова: Поддержка принятия решений; подразделение пограничного контроля; нейронная сеть; системный анализ.

CHOOSING A NEURAL NETWORK MODEL IN ORDER TO CREATE A DECISION SUPPORT SYSTEM

The article offers a variant of the neural network model for the development of a decision support system created for the head of the border control unit. The proposed model is based on a knowledge-based decision support system. Approaches to its implementation include scientific methods of system analysis and machine learning technology.

Keywords: Decision support; border control unit; neural network; system analysis.

Введение

В целях реализации основных задач органов пограничной службы прогнозирование и оценка обстановки на Государственной границе Республики Беларусь становятся весьма затруднительными ввиду растущего потока разнородной неоднозначной информации, характеризующейся современными военно-политическими и социально-экономическими преобразованиями, нарастанием противодействия интеграционным процессам нашего государства, а также возрастанием угрозы возникновения локальных и региональных конфликтов вблизи Государственной границы. В данных условиях становится весьма актуальным проведение научных исследований и принятием организационных мер, связанных с внедрением в служебную деятельность начальников подразделения пограничного контроля системы поддержки принятия решения (далее – СППР), позволяющую оказывать действенную помощь в ходе оценки обстановки, ее прогнозировании и в целом при принятии управленческого решения.

1. Методология исследования / теоретические основы

Исходя из проведенного анализа, на сегодняшний день весьма перспективными являются СППР, ориентированные на знания (Knowledge-Driven DSS). В основе таких систем заложены знания для машинного обучения, полученные на основе экспертных данных. Для разработки СППР, ориентированной на знания, возможно прибегнуть к одному из следующих методов:

1. Метод системного подхода, принципы которого реализуются на научной базе системного анализа, позволяющего разложить слабоструктурированную задачу на элементы и построить иерархию целей. Системный анализ является прикладной наукой, предназначенной для подготовки и обоснования решений по сложным слабоструктурированным проблемам в управленческой деятельности различных отраслей. Данному методу посвящено достаточное количество научных трудов. Наиболее известные его исследователи: А.А. Богданов, В.М. Глушков, С.А. Кузьмин, Ф.И. Перегудов, Ф.П. Тарасенко, В.Г. Афанасьев, Л.А. Петрушенко, В.А. Карташов, В. Н. Садовский, Ю. Г. Марков, И. В. Блауберг, Т. Гоббс, В.С. Тюхтин, М. Месарович, О. Конт, М. И. Сетров, Т. Саати, Д. Пауэр, Я. Такахара, В. Дхар, Р. Стайн и др.

Для решения задачи, стоящей перед нами, системный анализ был применен для декомпозиции факторов, влияющих на изменение обстановки в автодорожных пунктах пропуска, на частные показатели, как наиболее приверженные к быстрому изменению обстановки [1].

2. Метод машинного обучения для задач классификации, позволяющий оптимизировать иерархию целей для нивелирования большой размерности. Для решения данной задачи был выбран один из ансамблевых методов – классификатор экстремально рандомизированных деревьев (Extra Trees Classifier), который позволяет осуществить отбор признаков на основе их важности, что дало возможность исключить косвеннополезные показатели факторов, влияющие на изменение обстановки [2]. Данную отрасль науки развивают – Дж.Вандер Плас, А. Мюллер, С. Гвидо, Дж. Грас.

3. Метод анализа иерархий, как один из наиболее подходящих среди методов теории принятия решений с целью получения обоснованных количественных данных и сопоставления их с качественными. При решении данной задачи метод использовался в условиях неопределенности. Он позволяет учитывать многокритериальность и неопределенность, а также осуществить выбор решений из множеств альтернатив различного типа при наличии критериев, имеющих разные типы шкал измерения [3]. Наи-

более выдающимися ученными в данной области науки являются – А.В. Андрейчиков, О.Н. Андрейчикова, О.И. Ларичев, В.В. Подиновский, В.Д. Ногин, Т. Саати, Дж. фон Нейман, О. Моргенштерн, Л. Заде и др.

С применением данного метода в ходе исследования был проведен экспертный опрос и рассчитаны весовые коэффициенты показателей факторов, влияющих на изменение обстановки на Государственной границе.

4. Один из методов реализации искусственного интеллекта – нейронная сеть. В нашем случае использование многослойного персептрона, который относится к одному из классов нейронных сетей прямого распространения [4]. Нейронные сети прямого распространения широко используются и могут решать задачи прогнозирования и кластеризации. Разработке данной отрасли науки посвящены труды следующих ученых: Х. Саймон, Т. Рашид, У. Мак-Каллок, Ф. Розенблатт, Р.Каллан и др.

Основополагающим элементом разрабатываемой СППР для начальника подразделения пограничного контроля является нейронная сеть, алгоритм ее работы следующий (рисунок 1):

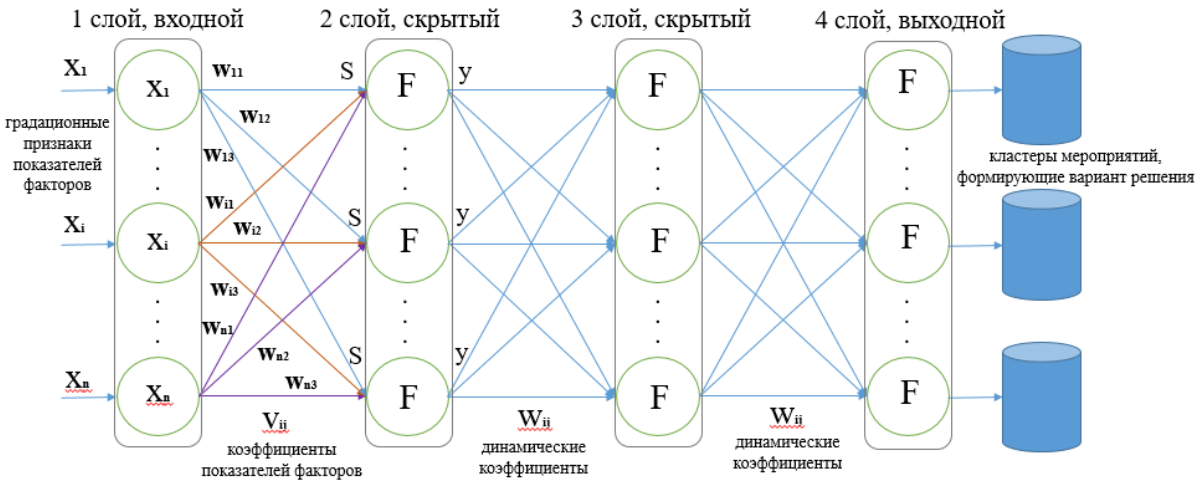


Рисунок 1 – Структурная модель нейронной сети системы поддержки принятия решения для начальника подразделения пограничного контроля

На входной слой нейронов поступают градационные признаки показателей факторов, которые представляют собой значения от 0 до 1, что в частности определяет влияние показателя фактора на итоговый вариант решения, который в итоге представляет собой комплекс мероприятий из баз знаний (баз данных). Такой подход позволяет распределить мероприятия, исходя из степени влияния показателей факторов на обстановку в пункте пропуска, а также повысить адекватность и точность оценки обстановки с последующим принятием управленческого решения.

Далее информация передается с помощью связей следующему слою, где каждая связь имеет собственный коэффициент веса, полученный с ис-

пользованием метода анализа иерархий, а следующий нейрон имеет входящие связи от каждого предыдущего слоя нейронов. Данные, полученные следующим нейроном, являются суммой всех данных от нейронов предыдущего слоя, перемноженных на коэффициенты весов (каждый на свой) [5].

$$S = \sum_{i=1}^{n} w_i x_i, \quad (1)$$

где S – взвешенная сумма входных сигналов, w_i – веса связей, x_i – поступающие значения градационных признаков показателей факторов.

Полученное значение сглаживается с помощью функции активации, в результате чего происходит формирование выходной информации. Вид функции активации может иметь различное выражение, выбор которого определяется характером решаемых задач [6].

Например, линейная функция активации:

$$y = k \cdot S. \quad (2)$$

Или пороговая бинарная функция активации:

$$y = sign(S) = \begin{cases} 1, S > 0; \\ 0, S \leq 0. \end{cases} \quad (3)$$

Линейная ограниченная функция активации:

$$y = \begin{cases} p, S > \alpha; \\ -p, S < -\alpha; \\ S, -\alpha \leq S \geq \alpha. \end{cases} \quad (4)$$

Сигмоидная функция активации:

$$y = 1/(1 + e^{-S}). \quad (5)$$

Это не полный перечень возможных функций активации, среди которых одной из широко применяемых в многослойных персептронах является сигмоидная функция, обладающая рядом преимуществ [7, 8, 9]:

усиление слабых сигналов и способность сопротивления к «насыщению» от мощных воздействий;

минимальная вычислительная нагрузка, что обеспечивается простым выражением для ее производной, при этом минимизируя вычислительную сложность метода обратного распространения ошибки;

наличие сглаженности в переходной области, в отличие от единичной ступенчатой функции, что обеспечивает непрерывность функции и, следовательно, позволяет ее дифференцировать;

монотонность и дифференцируемость на всей оси абсцисс.

Вместе с тем сигмоидная функция становится менее подходящей для многослойных персептронов, в которых применяется большое количество

скрытых слоев, что является причиной возникновения проблемы исчезающего градиента.

2. Результаты и их обсуждение

В решаемой задаче проблема исчезающего градиента исключена по причине малого числа слоев персептрона – входной, два скрытых и выходной слои [10]. Информация передается дальше до тех пор, пока не достигнет выхода (выходного слоя). В результате полученные выходные данные формируют комплекс мероприятий, необходимых для проведения начальником подразделения пограничного контроля, исходя из оцененной обстановки.

Стоит обратить внимание на то, что связи между вторым и третьим скрытыми слоями, а также третьим скрытым и четвертым выходным слоями имеют динамические коэффициенты, которые определяются по мере обучения нейронной сети посредством метода обратного распространения ошибки известного как метод градиентного спуска [5].

Суть данного метода заключается в получении наименьшего значения ошибки в выходных данных по отношению к тренировочным, таким образом получая точность прогноза посредством использования тренировочных данных при настройке параметров сети.

Наиболее часто используемой функцией для расчета ошибки является функция потерь среднеквадратической ошибки (MSE):

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2 \qquad (6)$$

Смысл заключается в нахождении квадратной суммы расстояния между значением тренировочных данных (целевое значение) и расчетным значением (фактическое значение на выходе нейронной сети) и усреднением его по всему набору данных, где N – количество экземпляров тренировочных данных [11].

С эксплуатационной стороны процесс обучения и работы нейронной сети в составе ССПР будет как на рисунке 2.

Начальник подразделения пограничного контроля использует значения градационных признаков и мероприятия, определяющие управленческое решение и хранящиеся в базе данных для их сопоставления, формируя при этом тренировочные данные.

СППР обрабатывает полученные тренировочные данные и подает их на вход нейронной сети, после чего входные данные проходят обработку внутри нейронной сети, формируя выходной сигнал и выводя комплекс мероприятий для начальника подразделения пограничного контроля.

Начальник подразделения пограничного контроля дает оценку прогноза и если прогноз неудовлетворительный, что указывает на недостаточность количества тренировочных экземпляров, то производится процесс переобучения сети путем создания дополнительных тренировочных данных. Процесс обучения продолжается до тех пор, пока ошибка не достигнет минимума, что свидетельствует о том, что нейронная сеть обучена.

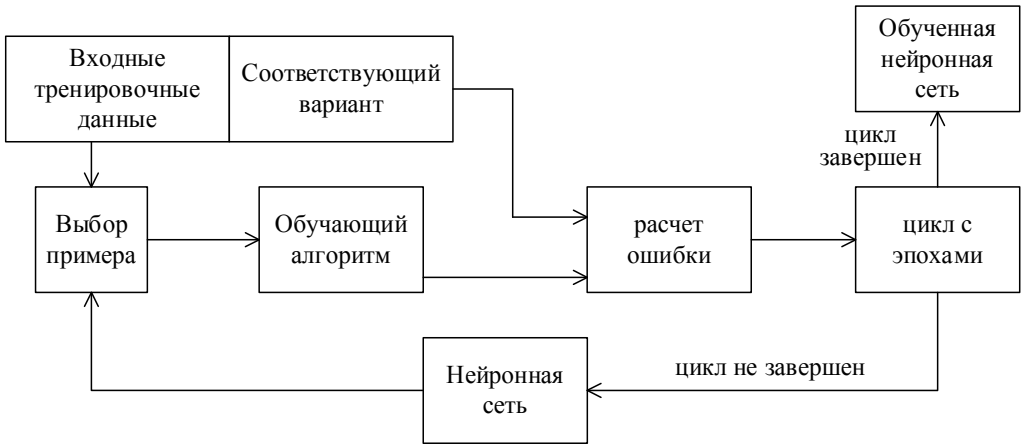


Рисунок 2 – Процесс работы и обучения нейронной сети СППР для начальника подразделения пограничного контроля

По полученным результатам представляется возможным определить следующие элементы ССПР для начальника подразделения пограничного контроля:

модуль сбора данных для обучения нейронной сети;
модуль обучения нейронной сети;
модуль прогнозирования нейронной сети;
модуль работы с данными;
интерфейс.

Заключение

Таким образом, ССПР для начальника подразделений пограничного контроля разрабатывается на основе многослойного персептрона, входными данными которого являются градационные признаки показателей факторов, сглаживающей функцией избрана сигмоида, а для расчета ошибки избрана функция потерь среднеквадратической ошибки. Для достижения точного результата прогнозирования выбрано два скрытых слоя, что позволяет избежать проблему исчезающего градиента, но при этом добиться адекватных прогнозирующих результатов. Обучение нейронной сети основано на методе с учителем на основе множества тренировочных

данных, представляющих собой пары «известный вход – известный выход», где входными данными являются градационные признаки показателей факторов, а выходами мероприятия характеризующие варианты решений, которые сопоставляются начальником подразделения пограничного контроля для формирования тренировочных данных и получения на выходе целевых значений.

Библиографические ссылки

1. Сорокин М.Н., Рябенко Д.С. Подходы к применению метода анализа иерархий в целях принятия решения начальником подразделения пограничного контроля // Системный анализ и прикладная информатика. 2021. № 3. С. 4–13.
2. Сорокин М.Н., Рябенко Д.С. Подходы к оптимизации модели принятия решения на охрану Государственной границы в пункте пропуска. Актуальные аспекты совершенствования пограничной безопасности: материалы международной научно-практической конференции Республиканского государственного учреждения «Пограничная академия Комитета национальной безопасности Республики Казахстан». Алматы: РГУ «ПА КНБ РК», 2021. С. 108–113.
3. Андрейчиков А.В., Андрейчикова О.Н. Анализ, синтез, планирование решений в экономике М.: Финансы и статистика, 2000. 368 с.
4. Каллан Р. Нейронные сети. Краткий справочник М.: Издательский дом «Вильямс», 2017. 288 с.
5. Тарик Р. Создаем нейронную сеть. СПб.: ООО «Диалектика», 2019. 272 с.
6. Хайкин С. Нейронные сети: полный курс М.: Издательский дом «Вильямс», 2006. 1104 с.
7. Черняк Е. Введение в глубокое обучение. СПб. : ООО «Диалектика», 2020. 192 с.
8. Шитиков В.К., Розенберг Г.С., Зинченко Т.Д. Количественная гидроэкология: методы системной идентификации. Тольятти: ИЭВБ РАН, 2003. 463 с.
9. The sigmoid activation function: activation in multilayer perceptron neural networks. All about circuits. URL: https://www.allaboutcircuits.com/technical-articles/sigmoid-activation-function-activation-in-multilayer-perceptron-neural-network/.
10. McClelland J. Explorations in parallel distributed processing: a handbook of models, programs, and exercises. Stanford, 2015. 249 p.
11. Мюллер А. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными. СПб.: ООО «Альфа-книга», 2018. 480 с.

УДК 550.8.05

«СЫРЫЕ» ДАННЫЕ И НЕКОТОРЫЕ РЕЦЕПТЫ ИХ «ПРИГОТОВЛЕНИЯ»

М.М. Татур[1], В.М. Проровский[1], Д.В. Куприянова[1], И.Н.Носырев[2]

[1]*Белорусский государственный университет информатики и радиоэлектроники, ул. П. Бровки, 6, г. Минск, 220013, Беларусь,*
tatur@bsuir.by, slawapro@gmail.com, diankupriyanova@gmail.com
[2]*Объединенный институт машиностроения Национальной академии наук Беларуси, ул. Академическая, 12, г. Минск, 220072, Беларусь,* nosureviluha@mail.ru

Подготовка данных для их обработки формальными алгоритмами анализа (классификации, кластеризации, регрессии и др.) имеет важное значение, поскольку существенно влияет на результат принимаемых решений. В работе рассматриваются типовые операции препроцессинга данных на примере набора данных о пожарах. Наряду с тривиальными операциями по очистке и форматированию, этап подготовки данных включает неформальные процедуры, которые требуют участия как специалистов по анализу данных, так и экспертов из предметной области. Показывается, как некоторые признаки необходимо преобразовывать из разряда порядковых, номинальных, необрабатываемых, в числовые значения, а также придавать вес в принятии решений.

***Ключевые слова*:** интеллектуальный анализ; машинное обучение; подготовка данных; статистические данные о пожарах.

«RAW» DATA AND SOME RECIPES FOR THEIR «COOKING»

M.M. Tatur[a], V.M. Prorovsky[a], D.V. Kupriyanova[a], I.N. Nosarau[b]

[a]*Belarusian State University of Informatics and Radioelectronics, P.Brouki st., 6, Minsk, 220013, Belarus,*
tatur@bsuir.by, slawapro@gmail.com, diankupriyanova@gmail.com
[b]*Joint Institute of Mechanical Engineering of the National Academy of Sciences of Belarus, Akademicheskaya str., 12, Minsk, 220072, Belarus,* nosureviluha@mail.ru

Preparation of data for their processing by formal analysis algorithms (classification, clustering, regression, etc.) is important, since it significantly affects the result of decisions made. The paper considers typical data preprocessing operations using the fire data set as an example. Along with the trivial cleansing and formatting, the data preparation phase includes informal procedures that require the participation of both data scientists and subject matter experts. It is shown how some signs need to be converted from the category of ordinal, nominal, unprocessed, into numerical values, and also to give weight in decision making.

Keywords: data mining; machine learning; data preparation; fire statistics.

Введение

В большинстве научных статей, посвященных интеллектуальному анализу данных и машинному обучению, рассматриваются результаты обработки тестовых или полученных (измеренных) наборов данных различными *формальными* алгоритмами. Как правило, предметом исследования выступает оценка эффективности примененных методов и обоснование выбора наиболее наилучшего метода, алгоритма и их параметров [1-4]. Не преуменьшая указанной стороны анализа данных, следует обращать внимание на этап подготовки данных (Data preparation в методологии CRISP-DM) [5], в котором значительная часть является *неформальной.*

В академической среде имеется устоявшееся мнение, что этап подготовки данных занимает примерно 60-80% времени и ресурсов, отведенных на анализ. В этой связи, возникает ряд вопросов, из которых назовем лишь три, на наш взгляд, наиболее значимые.

1. Существует ли общепринятая методика препроцессинга исходных (сырых) данных?
2. Какие этапы подготовки следует выделять и степени их формализации?
3. Насколько качество подготовки данных способно повлиять на конечные результаты анализа и, в конечном итоге, на принятие решений?

По первому вопросу, можно привести ряд публикаций [6-9], из которых следует, что общепринятой методики подготовки данных не существует. Даже при использовании современных библиотек алгоритмов автоматического машинного обучения (AutoML), осуществляющих предварительную обработку сырых данных этот процесс реализуется в каждом случае по-своему [10]. При ручной обработке дата-аналитик решает на основании своего опыта какие методы и в каком объеме использовать.

Второй вопрос будет более детально рассмотрен в настоящей работе.

Ответ на третий вопрос связан с проведением специальных исследований, детальным анализом результатов экспериментов и выходит за рамки данной публикации.

1. Пример исходных данных для обработки

Ниже, в таблице 1, приведен пример исходных данных о пожарах, содержащий признаки и данные, сходные с реальными.

Таблица 1 - Образец исходных данных о пожарах

	Дата	Адрес (текст)	Причина (текст)	Площадь пожара (число)	Объект пожара (текст)	Расстояние (число)	Собственник объекта (текст)
	1	2	3	4	5	6	7
1	15.01.2021	ул. Лиственная, 2/58	неосторожное обращение с огнем	43	жилой дом	15	Иванов В.В.
2	24.03.2021	ул. Шмидта, 5	поджог	60	склад	7,5	ООО “Пистон”
3	13.05.2021	дддлоолол		ошибка	жилой дом	-	Петров И.И.
4	18/07/2021	пр-д. Жемчуж-ный, 2в	неосторожное обращение с огнем	13	склад		УП “Промстрой”
5	01/09/2021		поджог	85	жилой дом	12	Сидоров С.Г.
…							
n	24.03.2021	ул. Шмидта, 5	поджог	60	склад	7,5	ООО “Пистон”

Число записей (строк) в данных исчисляется тысячами и даже десятками тысяч, в зависимости от выбранного для анализа исторического периода.

Число информативных признаков (столбцов) в данном примере для наглядности ограничено семью, в то время как реальные данные содержат десятки, детализирующие способы реагирования, причиненный ущерб и другую информацию о пожаре.

Несложно заметить, что данные в таблице имеют как числовое (площадь, расстояние), так порядковое (дата), так и номинальные (адрес, причина, объект, собственник) значения.

Заметим, что на текущий момент изложения, задача анализа данных не сформулирована.

Для начала, перечислим операции, которые обычно относят к этапу подготовки данных, с кратким пояснением их сути.

Эти операции включают:

- Очистку данных – выявление и исправление ошибок или ошибок в данных.

- Отбор признаков – определение признаков, которые наиболее важны для анализа.
- Преобразование данных – изменение масштаба или распределения переменных.
- Конструирование признаков – получение новых переменных из доступных данных.
- Уменьшение размерности – создание компактных проекций данных.

Данный перечень не претендует на полноту, а названия операций и их содержание также могут отличаться в изложении различных авторов.

2. Операции по предварительной обработке данных

2.1 Очистка данных

Включает исправление систематических проблем или ошибок в беспорядочных данных. Наиболее полезная очистка данных требует глубоких знаний предметной области и может включать выявление и устранение конкретных ошибочных наблюдений. Имеется много причин, по которым в данных могут быть неверные значения. Например, опечатки, повреждения, дублирование и т. д. Изучение предметной области может позволить идентифицировать явно ошибочные наблюдения, поскольку они отличаются от типовых. Например, отрицательное расстояние от пожарной части до места пожара.

После выявления беспорядочных, зашумленных, искаженных или ошибочных наблюдений их можно устранить. Это может быть удаление строк или столбцов, или замена наблюдений новыми значениями.

К общим операциям по очистке данных, можно отнести:

- Применение статистических методов для определения нормальности данных и выявления выбросов.
- Выявление дубликатов строк данных и их удаление
- Замена пустых значений как отсутствующих
- Замена пропущенных значений с использованием статистики или обученных моделей.

В таблице 1 строки 2 и n являются идентичными, и для дальнейшей работы одна из них должна быть удалена. В колонке 3 имеются пропуски, а в колонках 4 и 6 имеются не только пропуски, но и текстовые значения, которые не могут быть преобразованы в числа напрямую. В столбце 1 даты представлены в разных форматах. В качестве примера выявления выбросов в данных о пожарах можно привести реальное применение методов S-ESD (Seasonal Extreme Studentized Deviate) при построении прогнозной модели возникновения пожаров [11]. Кроме этого, исследование

выбросов в контексте интеллектуального анализа данных может выявить полезную, неизвестную ранее информацию.

Очистка данных это операция, которая обычно выполняется в первую очередь, перед другими операциями подготовки данных

2.2 Отбор признаков

Заключается в выборе подмножества входных признаков, наиболее релевантных прогнозируемой целевой переменной. Нерелевантные и избыточные входные переменные могут отвлекать или вводить в заблуждение алгоритмы обучения, что может привести к снижению эффективности прогнозирования. Более эффективно разрабатывать те модели, которые используют меньше признаков, необходимых для прогнозирования, т.е. наиболее простые из работающих достаточно точно моделей.

Методы выбора признаков условно подразделяются на использующие целевую переменную (supervised), и не использующие (unsupervised). Контролируемые методы могут быть далее разделены на модели, которые автоматически выбирают признаки в процессе подбора модели (встроенные), те, которые явно выбирают признаки, приводящие к наилучшей производительности модели (обертки), и те, которые оценивают каждую входную переменную и позволяют выбрать их несколько из набора (фильтры).

Распространено использование статистических операций (например, корреляция) для оценки входных признаков. Далее признаки ранжируются по их оценкам и набор с наилучшими показателями используется в качестве входных данных для модели. Выбор статистической операции зависит от типа данных входных переменных. Различают общие варианты выбора признаков, например:

- Категориальные входные данные для целевой переменной классификации.
- Числовые входные данные для целевой переменной классификации.
- Числовые входные данные для целевой переменной регрессии.

В случае, если типы данных входных переменных отличаются, используют разные методы фильтрации. Как альтернативу можно использовать метод-оболочку, например, метод устранения рекурсивных признаков (RFE), который не зависит от типа входной переменной.

В нашем примере очевидным для исключения столбцами является 7. А вот данные из столбца 2, несмотря на имеющиеся ошибки могут оказаться полезными. Используя вспомогательные сервисы вроде Nominatim от OpenStreetMap возможно получение для дальнейшего применения геокоординат по текстовой части информации.

2.3 Преобразование данных.

Эти операции используются для изменения типа переменных или распределения данных. Они представляют собой достаточно большой набор различных методов, применяемых как к входным, так и к выходным переменным. Данные могут иметь один из нескольких типов:

Числовые	Категориальные
Integer – целые числа без дробной части. Float – числа с плавающей запятой.	Ordinal – метки с ранговым порядком. Nominal – метки без ранжирования. Boolean – значения True и False.

При необходимости возможно преобразовать числовую переменную в порядковую переменную (дискретизация). Для большинства задач классификации категориальные переменные могут быть закодированы в виде целого числа или логических переменных.

- Дискретизация – кодирование числовой переменной как порядковой переменной.
- Порядковое преобразование – кодирование категориальной переменной в целочисленную переменную.
- Быстрое кодирование (One-Hot encoding) – кодирование категориальной переменной в двоичные переменные.

Масштабирование числовых данных, т.е. их обезличивание, в отношении единиц измерения. Операция формализуемая или почти формализуемая.

Существует ряд способов масштабирования, среди которых наиболее часто применяются такие как нормализация и стандартизация. Нормализация – это преобразование значений отдельного признака в диапазон значений с плавающей точкой от 0 до 1. Стандартизация преобразовывает значения признака путем вычитания среднего значения (так называемое центрирование) и деления на стандартное отклонение, чтобы сдвинуть распределение так, чтобы среднее значение равнялось нулю, а стандартное отклонение равнялось единице.

Для некоторых алгоритмов масштаб числовых значений признака не влияет. В первую очередь это деревья решений и ансамбли деревьев, такие как случайный лес.

Выбор необходимости и способа масштабирования определяется дата-аналитиком, исходя из конкретной решаемой задачи.

Распределение значений вероятности для числовых переменных может быть изменено. Если распределение почти нормальное, но искажено или смещено, его можно сделать более нормальным с помощью степенно-

го преобразования [12]. Квантильные преобразования можно использовать для принудительного распределения вероятностей (равномерной или нормальной) для переменной с необычным естественным распределением.

При преобразовании данных операции обычно выполняются отдельно для каждой переменной.

В нашем наборе данных столбцы 3 и 5 являются категориальными признаками и с помощью порядкового преобразования или быстрого кодирования могут быть преобразованы к формату, который далее может быть использован для обучения конкретных моделей.

2.4 Конструирование признаков.

Процесс создания новых входных переменных из доступных данных называется конструированием признаков. Эти операции сильно зависят от состава и типов данных. При этом необходимо сотрудничество эксперта в предметной области. Эти условия затрудняют унификацию общих методов. Есть ряд приемов, которые можно использовать повторно:

- Добавление логической переменной флага для некоторого состояния.
- Добавление групповой или глобальной сводной статистики, например, среднего значения.
- Добавление новых переменных для каждого компонента составной переменной, такой как дата-время.

Например, при построении обобщенной линейной модели обстановки с пожарами [13] выполнено разложение признака с типом «дата» на 21 признак, значения весов которых приведены в таблице 2.

Таблица 2 – Признаки и их веса

№	Показатель	Вес	№	Показатель	Вес
1	Разность дат (Текущая дата - Дата)	0,726	8	Дата: полугодие = 2	0,085
2	Дата: день недели = 7	0,509	9	Дата: месяц года	0,085
3	Дата: день недели = 1	0,436	10	Дата: полугодие = 1	0,030
4	Дата: день месяца	0,245	11	Дата: день недели = 6	0,022
5	Дата: месяц квартала = 1	0,182	12	Дата: день недели = 5	0,018
6	Дата: день недели = 2	0,168	13	Остальные показатели	0,000
7	Дата: квартал = 2	0,108			

Подход, основанный на статистике, заключается в создании копий числовых входных переменных, которые были изменены с помощью простой математической операции, такой как возведение их в степень или умножение на другие входные переменные, называемые полиномиальными признаками.

Полиномиальное преобразование заключается в создании копий числовых входных переменных, возведенных в степень.

Основная цель конструирования признаков состоит в том, чтобы добавить более широкий контекст к отдельному наблюдению или разложить сложную переменную для более прозрачного представления о входных данных. Конструирование признаков некоторые авторы относят к разновидности преобразования данных.

2.5 Уменьшение размерности

Количество входных признаков для набора данных рассматривается как размерность данных. Например, две входные переменные образуют двумерную область, где каждая строка данных определяет точку в этом пространстве. Набор данных может содержать любое количество входных переменных для создания многомерного пространства. Проблема в том, что чем больше измерений имеет это пространство, тем больше вероятность того, что набор данных представляет собой избыточную и, вероятно, нерепрезентативную выборку.

Это заставляет дата-аналитика выбирать наиболее информативные признаки. Альтернативой выбору признаков является создание проекции данных в пространство более низкого измерения, которое по-прежнему сохраняет наиболее важные свойства исходных данных. Этот процесс называется уменьшением размерности. В отличие от выбора признаков, переменные в прогнозируемых данных не связаны напрямую с исходными входными переменными, что затрудняет интерпретацию прогноза. Наиболее распространенным подходом к уменьшению размерности является использование методов матричной факторизации: анализ главных компонент, сингулярное разложение.

Основное преимущество этих методов заключается в том, что они устраняют линейные зависимости между входными переменными, например, коррелированными переменными. [9].

Заключение

Перечень рассмотренных операций предварительной подготовки данных может быть расширен, например, такой процедурой как оценка репрезентативности выборки для анализа данных и принятия решений. Данный вопрос тесно связан с такими параметрами выборки как количество наблюдений, размерность пространства признаков, распределение объектов (образов) в этом пространстве, наличие шумов, выбросов и др. Т.е. оценка репрезентативности тесно связана с другими операциями подготовки данных и, в некоторых случая, может быть представлена их композицией.

Также надо принимать во внимание целевую задачу, решаемую в ходе предстоящего анализа данных, которая в свою очередь будет существенно определять состав и содержание операций предварительной обработки.

Библиографические ссылки

1. Borges L. Comparison of data mining techniques and tools for data classification / L. Borges, M.Viriato, B.Jorge // ACM International Conference Proceeding Series. 2013. P. 113–116., DOI:10.1145/2494444.2494451.
2. Tapak L. Real-data comparison of data mining methods in prediction of diabetes in Iran / L.Tapak, H.Mahjub, O.Hamidi, J.Poorolajal // Healthc Inform Res. 2013. Vol. 19(3). P. 177–185., DOI: 10.4258/hir.2013.19.3.177.
3. Garg S. Comparative Analysis of Data Mining Techniques on Educational Dataset / S. Garg, A. K. Sharma // International Journal of Computer Applications. 2013. Vol. 74(5). P. 0975 – 8887.
4. Trifonov R. Analysis of data mining evaluation methods' efficiency / R. Trifonov, D. Gotseva, V. Angelov // International Journal of Development Research. 2017. Vol 11(7). P. 16880–16884.
5. Azevedo A. KDD, SEMMA and CRISP-DM: A parallel overview [Electronic resource] / A. Azevedo, M. Santos // IADIS Multi Conf. on Computer Science and Information Systems, Amsterdam, 22–27 July 2008 / Intern. Assoc. for Development if the Inform. Soc.; Associate Ed.: Luís Rodrigues and Patrícia Barbosa. Amsterdam, 2008. P. 182–185.
6. Shichao Z. Data Preparation for Data Mining / Z. Shichao, Z. Chengqi, Y. Qiang. // Applied Artificial Intelligence. 2003. Vol. 17. P. 375–381., DOI:10.1080/713827180.
7. Zheng, A. Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists / A. Zheng, A.Casari. 1st Edition. O'Reilly Media, 2018. 218 p.
8. Kuhn M. Feature Engineering and Selection: A Practical Approach for Predictive Models / M.Kuhn, K.Johnson. 1st Edition. Chapman & Hall/CRC, 2019. 298 p.
9. Brownlee J. Data Preparation for Machine Learning. 1st Edition. 2020. 398 p.
10. Чистый AutoML для "грязных" данных: как и зачем автоматизировать предобработку таблиц в машинном обучении. URL: https://habr.com/ru/company/ods/blog/657525/. (дата обращения: 24.06.2022.)

11. Tatur M.M. Exploratory analysis of the fire statistics using automatic time series decomposition / M.M. Tatur, V.M. Prorovsky, A.G. Ivanitskiy, M. Kvassay // Information and Digital Technologies 2021: Proc. of the Intern. Conf., 22–24 June 2021, Zilina, Slovakia, ed. J. Rabcan [et al.]. Zilina: University of Zilina, 2021. P. 158–161.
12. Box G.E.P. An Analysis of Transformations / G.E.P. Box; D.R. Cox // Journal of the Royal Statistical Society. Series B (Methodological), Vol. 26, №. 2. 1964. P. 211–252.
13. Татур М.М. и др. Сравнение точности алгоритмов автоматического машинного обучения при прогнозировании обстановки с пожарами на объектах жилого сектора //Чрезвычайн. ситуации: предупреждение и ликвидация. 2021. № 22(50). С. 61–70.

УДК 519.24

МОМЕНТЫ ПЕРВЫХ ДВУХ ПОРЯДКОВ ОЦЕНКИ СЕМИВАРИОГРАММЫ СЛУЧАЙНОГО ПРОЦЕССА

Т.В. Цеховая, Д.А. Мармузевич

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, tsekhavaya@bsu.by, marmuzev@bsu.by

Для случайного процесса с дискретным временем найдены выражения для математического ожидания, дисперсии, ковариационной функции, семивариограммы и спектральной плотности, исследована стационарность процесса. Построена оценка семивариограммы рассматриваемого случайного процесса, получены выражения для первых двух моментов исследуемой статистики.

Ключевые слова: случайный процесс; семивариограмма; оценка; стационарность в широком смысле; внутренняя стационарность.

MOMENTS OF THE FIRST TWO ORDERS OF ESTIMATION OF THE SEMIVARIOGRAM OF A RANDOM PROCESS

T.V. Tsekhavaya, D.A. Marmuzevich

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus, tsekhavaya@bsu.by, marmuzev@bsu.by

For a random process with discrete time, expressions for the expected value, variance, covariance function, semivariogram, and spectral density are found, and the stationarity of the process is investigated. An estimate of the semivariogram of the random process under consideration is constructed, expressions for the first two moments of the statistics under study are obtained.

Keywords: random process; semivariogram; estimate; second-order stationarity; intrinsic stationarity.

Введение

В настоящее время для решения многих прикладных задач прогнозирования применяется геостатистический подход, в частности, метод кригинг. В основе кригинга лежит семивариограмма. В связи с этим актуальны задачи исследования свойств семивариограммы, а также построения и изучения оценок этой функции.

Результаты исследований свойств семивариограммы изложены, например, в работах [1–3]. Статистические свойства различных оценок се-

мивариограммы случайных процессов исследовались, например, в [1, 4-6]. В данной статье для гауссовского случайного процесса с дискретным временем найдены выражения для ковариационной функции, семивариограммы и спектральной плотности, также построена оценка семивариограммы рассматриваемого случайного процесса, получены выражения для ее первых двух моментов.

1. Теоретические основы

Рассмотрим случайный процесс $Z(t) = \sum_{i=1}^{p} \beta_i X_i(t)$, где $t \in Z,\ p \in N$, $\beta_i -$ константы, такие что:

$$\sum_{i=1}^{p} \beta_i^2 < \infty, \qquad (1)$$

а $X_i(t) -$ гауссовские стационарные случайные процессы с нулевым математическим ожиданием, ковариационными функциями $R_i(t),\ t \in Z$, спектральными плотностями $f_i(\lambda),\ \lambda \in \Pi = [-\pi; \pi]$.

Будем полагать, что взаимные ковариационные функции $R_{ij}(t_1, t_2)$, $t_1, t_2 \in Z$, случайных процессов $X_i(t)$ и $X_j(t),\ i, j = \overline{1, p},\ i \neq j$, удовлетворяют равенству: $R_{ij}(t_1, t_2) = M\left[X_i(t_1) X_j(t_2)\right] = 0$.

Исследуем процесс $Z(t)$ на стационарность. Для этого найдем его характеристики первых двух порядков во временной области. Легко показать, что $M\left(Z(t)\right) = 0$, ковариационная функция $R_z(t_1, t_2)$ имеет вид $R_z(t_1, t_2) = \sum_{i=1}^{p} \beta_i^2 R_i(t_1, t_2) = \sum_{i=1}^{p} \beta_i^2 R_i(t_1 - t_2) = R_z(t_1 - t_2)$, а дисперсия $DZ(t) = R_z(t, t) = R_z(0) = \sum_{i=1}^{p} \beta_i^2 R_i(0)$.

Учитывая условие (1), имеем $DZ(t) < \infty$. Таким образом, случайный процесс $Z(t)$ является стационарным в широком смысле в силу соответствующего определения.

Применяя связывающее соотношение между ковариационной функцией и семивариограммой стационарного в широком смысле случайного процесса [1], получим выражение для семивариограммы процесс $Z(t)$:

$$\gamma_z(t) = R_Z(0) - R_Z(t) = \sum_{i=1}^{p} \beta_i^2 \left(R_i(0) - R_i(t)\right), \qquad t \in Z.$$

Отсюда вытекает, что случайный процесс $Z(t)$ является также внутренне стационарным.

По определению спектральной плотности

$$f_Z(\lambda) = \frac{1}{2\pi} \sum_{t=-\infty}^{+\infty} R_Z(t) e^{-i\lambda t} = \frac{1}{2\pi} \sum_{t=-\infty}^{+\infty} \sum_{i=1}^{p} \beta_i^2 R_i(t)\, cos\lambda t,\ \ \lambda \in \Pi.$$

Следует отметить, что случайный процесс $Z(t)$ является гауссовским как линейная комбинация гауссовских случайных процессов.

Предположим далее, что $Z(1), \dots, Z(n) - n$ последовательных наблюдений за процессом $Z(t),\ t \in Z$. В качестве оценки семивариограммы рассмотрим статистику вида:

$$\widehat{\gamma_Z}(h) = \frac{1}{2(n-h)} \sum_{t=1}^{n-h} \big(Z(t) - Z(t+h)\big)^2, \qquad (2)$$

где $h = 0, \dots, n-1$. Также учтем, что $\hat{\gamma}(h) = \hat{\gamma}(-h),\ h = 0, \dots, n-1$ и $\hat{\gamma}(h) = 0$, где $|h| \geq n$.

Найдем выражения для первых двух моментов статистики (2) через временные и частотные характеристики процесса $Z(t)$.

2. Результаты

Теорема 1. *Для оценки $\widehat{\gamma_z}(h)$ имеют место следующие соотношения:*

$$M\widehat{\gamma_z}(h) = \gamma_z(h), \qquad (3)$$

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1), \hat{\gamma}_Z(h_2)\} = \frac{1}{2(n-h_1)(n-h_2)} \sum_{t_1=1}^{n-h_1} \sum_{t_2=1}^{n-h_2} \sum_{i=0}^{p} \sum_{j=0}^{p} \beta_i^2 \beta_j^2 (R_i(t_1-t_2)R_j(t_1-t_2) -$$
$$-2R_i(t_1-t_2)R_j(t_1-t_2-h_2) + R_i(t_1-t_2-h_2)R_j(t_1-t_2-h_2) - 2R_i(t_1-t_2)R_j(t_1+h_1-$$
$$-t_2) + 2R_i(t_1-t_2)R_j(t_1+h_1-t_2-h_2) + 2R_i(t_1-t_2-h_2)R_j(t_1+h_1-t_2) - 2R_i(t_1-t_2-$$
$$-h_2)R_j(t_1+h_1-t_2-h_2) + R_i(t_1+h_1-t_2)R_j(t_1+h_1-t_2) - 2R_i(t_1+h_1-t_2)R_j(t_1+h_1-$$
$$-t_2-h_2) + R_i(t_1+h_1-t_2-h_2)R_j(t_1+h_1-t_2-h_2). \qquad (4)$$

$$D\hat{\gamma}_Z(h) = \frac{1}{(n-h)^2} \sum_{t_1=1}^{n-h_1} \sum_{t_2=1}^{n-h_2} \sum_{i=1}^{p} \sum_{j=1}^{p} \beta_i^2 \beta_j^2 (4R_i(t_1-t_2)R_j(t_1-t_2) - 2R_i(t_1-t_2)R_j(t_1-$$
$$-t_2-h) + R_i(t_1-t_2-h)R_j(t_1-t_2-h) - 2R_i(t_1-t_2)R_j(t_1+h-t_2) + 2R_i(t_1-t_2-h)\times$$
$$\times R_j(t_1+h-t_2) - 2R_i(t_1-t_2-h)R_j(t_1-t_2) + R_i(t_1+h-t_2)R_j(t_1+h-t_2) -$$
$$-2R_i(t_1+h-t_2)R_j(t_1-t_2)),$$

$h_1, h_2, h = 0, 1, \dots, n-1, R_j(t) -$ ковариационные функции процессов $X_j(t),\ t \in Z, j = 1, \dots, p$.

Доказательство. Из определения семивариограммы и свойств математического ожидания, утверждение (3) теоремы вытекает очевидным образом.

Воспользуемся определением ковариационной функции, выражением для оценки семивариограммы (2) и свойствами математического ожидания, применим элементарные преобразования. Тогда

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\frac{1}{4(n-h_1)(n-h_2)}\sum_{t_1=1}^{n-h_1}\sum_{t_2=1}^{n-h_2}M[Z^2(t_1)Z^2(t_2)]-2M[Z^2(t_1)Z(t_2)\times$$
$$\times Z(t_2+h_2)]+M[Z^2(t_1)Z^2(t_2+h_2)]-2M[Z(t_1)Z(t_1+h_1)Z^2(t_2)]+4M[Z(t_1)\times$$
$$\times Z(t_1+h_1)Z(t_2)Z(t_2+h_2)]-2M[Z(t_1)Z(t_1+h_1)Z^2(t_2+h_2)]+M[Z^2(t_1+h_1)Z^2(t_2)]-$$
$$-2M[Z^2(t_1+h_1)Z(t_2)Z(t_2+h_2)]+M[Z^2(t_1+h_1)Z^2(t_2+h_2)]-(M[Z^2(t_1)]M[Z^2(t_2)]-$$
$$-2M[Z^2(t_1)]M[Z(t_2)Z(t_2+h_2)]+M[Z^2(t_1)]M[Z^2(t_2+h_2)]-2M[Z(t_1)Z(t_1+h_1)]\times$$
$$\times M[Z^2(t_2)]+4M[Z(t_1)Z(t_1+h_1)]M[Z(t_2)Z(t_2+h_2)]-2M[Z(t_1)Z(t_1+h_1)]\times$$
$$\times M[Z^2(t_2+h_2)]+M[Z^2(t_1+h_1)]M[Z^2(t_2)]-2M[Z^2(t_1+h_1)M[Z(t_2)Z(t_2+h_2)]+$$
$$+M[Z^2(t_1+h_1)]M[Z^2(t_2+h_2)]).$$

Из определения смешанного момента четвертого порядка, используя связывающее соотношение смешанных моментов со смешанными семиинвариантами, учитывая свойства процесса $Z(t)$, получим:

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\frac{1}{4(n-h_1)(n-h_2)}\sum_{t_1=1}^{n-h_1}\sum_{t_2=1}^{n-h_2}2R_Z(t_1,t_2)R_Z(t_1,t_2)-4R_Z(t_1,t_2)\times$$
$$\times R_Z(t_1,t_2+h_2)+2R_Z(t_1,t_2+h_2)R_Z(t_1,t_2+h_2)-4R_Z(t_1,t_2)R_Z(t_1+h_1,t_2)+$$
$$+4R_Z(t_1,t_2)R_Z(t_1+h_1,t_2+h_2)+4R_Z(t_1,t_2+h_2)R_Z(t_1+h_1,t_2)-4R_Z(t_1,t_2+h_2)\times$$
$$\times R_Z(t_1+h_1,t_2+h_2)+2R_Z(t_1+h_1,t_2)R_Z(t_1+h_1,t_2)-4R_Z(t_1+h_1,t_2)\times$$
$$\times R_Z(t_1+h_1,t_2+h_2)+2R_Z(t_1+h_1,t_2+h_2)R_Z(t_1+h_1,t_2+h_2).$$

Учитывая стационарность случайных процессов $X_i(t)$, $t\in Z$, получим требуемое равенство (4).

Отметим, что $D\hat{\gamma}_Z(h)=\operatorname{cov}\{\hat{\gamma}_Z(h),\hat{\gamma}_Z(h)\}$.

Найдем выражения для вторых моментов оценки семивариограммы через спектральные плотности процесса $Z(t)$.

Теорема 2. *Для ковариации и диспесии оценки семивариграммы, задаваемой равенством (2), справедливы соотношения соответственно:*

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2\Big(\frac{\pi}{n-h^-}\int_{\Pi}F_{ij}(x)\cos\frac{(h_1-h_2)x}{2}\Phi_{n-h^+}(x)dx+$$
$$+\frac{1}{2(n-h_1)(n-h_2)}\int_{\Pi}F_{ij}(x)\Delta_{h^+-h^-}(x)\Delta_{n-h^+}(x)\cos\frac{(n-h^+)x}{2}dx\Big),\qquad(5)$$

$$D\hat{\gamma}_Z(h)=\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2\frac{\pi}{n-h^-}\int_{\Pi}F_{ij}(x)\Phi_{n-h}(x)dx,\qquad(6)$$

$$F_{ij}(x)=\int_{\Pi} f_i(x-y)\,f_j(y)g(x,y)dy, \tag{7}$$

$f_i(\lambda), \lambda \in \Pi, -$спектральные плотности процессов $X_i(t),\ t \in Z,\ i=1,\dots,p,$

$$g(x,y)=(1-2e^{-iyh_2}+e^{-ixh_2}-2e^{iyh_1}+2e^{iy(h_1-h_2)}+2e^{-i(x-y)h_2+iyh_1}- \\ -2e^{iyh_1-ixh_2}+e^{ixh_1}-2e^{ixh_1-iyh_2}+e^{ix(h_1-h_2)})e^{i\frac{h_2-h_1}{2}x}, \tag{8}$$

$$\Phi_T(x)=(2\pi T)^{-1}\Delta_T^2(x)-\text{ядро Фейера}, \tag{9}$$

$$h^+=\max(h_1,h_2); h^-=\min(h_1,h_2), h, h_1, h_2=0,\dots,n-1, \tag{10}$$

$$\Delta_T(x)=\frac{\sin Tx/2}{\sin x/2}, T \in N, x \in R. \tag{11}$$

Доказательство. Рассмотрим (4). Используя связывающее соотношение между ковариационной функцией и спектральной плотностью стационарного в широком смысле процесса $Z(t)$, t ϵ Z, запишем:

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\frac{1}{2(n-h_1)(n-h_2)}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2[\iint_{\Pi^2} f_i(x_1)\,f_j(x_2)(1-2e^{-ix_2h_2}+ \\ +e^{-ih_2(x_1+x_2)}-2e^{ix_2h_1}+2e^{ix_2(h_1-h_2)}+2e^{--ix_1h_2+ix_2h_1}-2e^{ix_2h_1-ih_2(x_1+x_2)}+e^{i(x_1+x_2)h_1}- \\ -2e^{ih_1(x_1+x_2)-ix_2h_2}+e^{i(x_1+x_2)(h_1-h_2)})\sum_{t_1=1}^{n-h_1}e^{it_1(x_1+x_2)}\sum_{t_2=1}^{n-h_2}e^{-it_2(x_1+x_2)}dx_1dx_2].$$

Сделаем замену переменных интегрирования $x_1=x-y,\ x_2=y,$ учитывая элементарное соотношение $\sum_{t=1}^{T}e^{itx}=\Delta_T(x)e^{i\frac{T+1}{2}x}$, получим:

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\frac{1}{2(n-h_1)(n-h_2)}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2[\iint_{\Pi^2} f_i(x-y)\,f_j(y)(1-2e^{-iyh_2}+ \\ +e^{-ixh_2}-2e^{iyh_1}+2e^{iy(h_1-h_2)}+2e^{--i(x-y)h_2+iyh_1}-2e^{iyh_1-ixh_2}+e^{ixh_1}-2e^{ixh_1-iyh_2}+e^{ix(h_1-h_2)})\times \\ \times\Delta_{n-h_1}(x)e^{i\frac{n-h_1+1}{2}x}\Delta_{n-h_2}(x)e^{-i\frac{n-h_2+1}{2}x}d(x-y)dy]= \\ =\frac{1}{2(n-h_1)(n-h_2)}\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2[\iint_{\Pi^2} f_i(x-y)\,f_j(y)g(x,y)\Delta_{n-h_1}(x)\Delta_{n-h_2}(x)dxdy],$$

где функции $g(x,y), \Delta_T(x)$ имеют вид (8) и (11) соответственно.

Рассмотрим случай $h_1 > h_2$. Используем элементарное тригонометрическое равенство и соотношение (9), получим:

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2\Big(\frac{\pi}{n-h_2}\int_{\Pi}F_{ij}(x)\cos\frac{(h_1-h_2)}{2}\Phi_{n-h_1}(x)dx+ \\ +\frac{1}{2(n-h_1)(n-h_2)}\int_{\Pi}F_{ij}(x)\Delta_{h_1-h_2}(x)\Delta_{n-h_1}(x)\cos\frac{(n-h_1)x}{2}dx\Big), \tag{12}$$

где функция $F_{ij}(x)$ задается соотношением (7).

Аналогично запишем для случая $h_1 \le h_2$:

$$\operatorname{cov}\{\hat{\gamma}_Z(h_1),\hat{\gamma}_Z(h_2)\}=\sum_{i=1}^{p}\sum_{j=1}^{p}\beta_i^2\beta_j^2\Big(\frac{\pi}{n-h_1}\int_{\Pi}F_{ij}(x)\cos\frac{(h_2-h_1)}{2}\Phi_{n-h_2}(x)dx+ \\ +\frac{1}{2(n-h_1)(n-h_2)}\int_{\Pi}F_{ij}(x)\Delta_{h_2-h_1}(x)\Delta_{n-h_2}(x)\cos\frac{(n-h_2)x}{2}dx\Big). \tag{13}$$

Использовав соотношение (10) и объединив выражения (12), (13), получим выражение (5). Положив $h_1 = h_2 = h$ в равенстве (5), получим выражение (6).

Библиографические ссылки

1. Cressie N. Statistics for Spatial Data. New York: Wiley, 1991. 900 p.
2. Цеховая Т.В. Свойства вариограммы внутренне стационарных случайных процессов // Теория вероятностей, математическая статистика и их приложения. Материалы научной конференции. Минск, БГУ. 2004. С. 181–186.
3. Цеховая Т.В. Свойства внутренне стационарных случайных процессов // Журн. Белорус. гос. ун-та. Математика. Информатика, 2017. № 1. С. 28–33.
4. Цеховая Т.В. Асимптотическое распределение оценки семивариограммы гауссовского случайного процесса // Вестн.БГУ. Сер. 1, Физика. Математика. Информатика. 2015. № 1. С. 89 – 95.
5. Цеховая Т.В. Асимптотическое распределение оценки вариограммы // Вестник БрГУ им. А.С. Пушкина. 2008. №2(31). С. 32–37. Network //Mathematical Geosciences. 2022. Т. 54. №. 1. С. 177–205.
6. Mazzella A., Mazzella A. The importance of the model choice for experimental semivariogram modeling and its consequence in evaluation process // Journal of Engineering. 2013. 11 p.

УДК 519.168

ПОДХОДЫ К РЕШЕНИЮ ЗАДАЧИ О МНОГИХ СТАНКАХ

И.Г. Яр-Мухамедов

Институт машиноведения и автоматики НАН КР
ул. Скрябина, 23, 720055, г. Бишкек, Кыргызстан. aldar@email.su

Рассмотрены подходы к решению задачи о многих станках. Показана ситуация, когда невозможно применение известных правил и алгоритмов к решению задачи. Обоснована необходимость формулирования новых правил.

Ключевые слова: задача о многих станках; алгоритм Джонсона; правила запуска деталей в обработку.

APPROACHES TO RESOLVING THE MULTI-STAGE PRODUCTION SCHEDULES

I.G. Yar-Mukhamedov

The Institute of machine science and automation of the national Academy of Sciences of the Kyrgyz Republic, 23 Scryabin st., Bishkek, 720055, Kyrgyzstan. aldar@email.su

A collection of production items and some machines or stages are given. The setup plus work time is set for each item to pass through each stage. The subjects under consideration are: the scope of effective use of Johnson's decision rule and its generalization; the need for qualitatively new approaches to solving the problem

Keywords: multi-stage scheduling problem; Johnson's rule; alternative approach.

Введение

Задача, сформулированная Р. Беллманом и решенная для некоторых случаев С. Джонсоном [1] состоит в том, что для партии деталей, подлежащих обработке на станочной линии, заданы времена обработки, включая подготовительно-заключительное время, и требуется определить порядок запуска деталей в обработку таким образом, чтобы суммарное время обработки всей партии было минимальным. С. Джонсон разработал правило и алгоритм, которые позволяли находить оптимальные решения для линии из двух станков, дал им аналитическое обоснование. Для линии из трех станков общего решения найти не удалось и он ограничился одним частным случаем.

Эта, казалось бы, простая задача привлекала внимание многих исследователей. Предлагались различные эвристические правила и алгоритмы.

Но более или менее общего и теоретически обоснованного метода пока нет.

1. Методология исследования / теоретические основы

Декомпозиционный подход в решении задачи о многих станках состоит в разбиении исходной задачи на ряд задач меньшей размерности и применении к ним алгоритмов, сходных с алгоритмом С. Джонсона. Недостаток этого подхода в непроработанности, отсутствии критериев разбиения и выбора правил для применения на каждом из этапов решения задачи. Попытки в этом направлении наиболее многочисленны, но отсутствие значимых результатов не позволяет останавливаться на них более подробно.

Редуционисткий подход заключается в преобразовании исходной задачи в задачу о двух станках [2]. Преобразование осуществляется над исходными данными и заключается в вычислении суммарных времен ожидания начала обработки и суммарных времен ожидания окончания обработки для каждой из деталей на всей станочной линии. В наглядной форме преобразование может быть представлено в виде таблицы 1.

Таблица 1 – Иллюстрация результатов свертки данных

№№ деталей	Данные исходной задачи о многих станках, указаны времена обработки на станках					→	Результат редукции	
	1	2	3	4	5		ВОН	ВОК
1	4	3	6	6	2		40	40
2	3	4	3	2	4		32	32
3	1	6	3	2	3		30	30
4	4	1	2	3	3		26	26
5	6	1	4	1	6		36	36
6	1	3	4	1	2		22	22
7	5	4	1	4	5		38	38

Сокращения ВОН и ВОК использованы для обозначения суммарных времен ожидания начала и окончания обработки деталей. В таблице размещены данные о деталях, подлежащих последовательной обработке на пяти станках. Число деталей в партии равно семи. В данном примере времена обработки являются равномерно распределенными случайными величинами на интервале от единицы до шести.

К редуцированным данным могут быть применены правила и алгоритм С. Джонсона. При этом вместо времени обработки на первом станке рассматривается показатель ВОН, а времени обработки на втором – ВОК.

Вместе с тем анализ особенностей задачи о многих станках показывает, что известные правила упорядочения деталей не могут быть, в об-

щем случае, универсальными и эффективными. Если станков в линии больше двух, становится значимым фактор взаимовлияний времен обработки деталей, приводящий к изменению значения критериального показателя. Причем эти взаимовлияния не обусловлены, по крайней мере непосредственно, соотношениями времен ожидания начала и окончания обработки. В таблице 1 как раз и представлен пример, для которого времена ожиданий для деталей равны и правила С. Джонсона неприменимы.

Не смотря на равенство времен ожидания наблюдается явная дифференциация решений по степени оптимальности. На рисунке представлено распределение количества решений в зависимости от значения критерия.

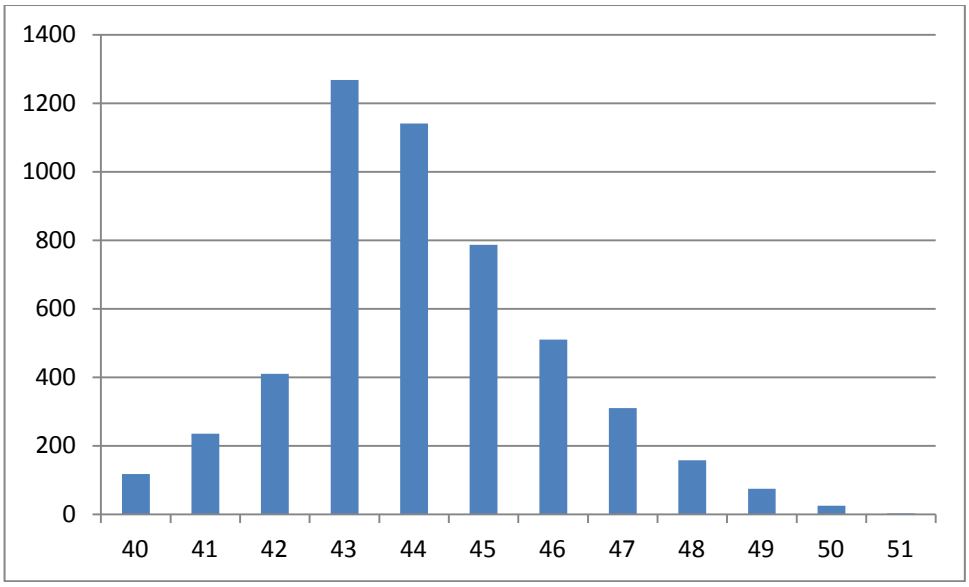


Рисунок – Частоты встречаемости решений (по вертикальной оси) в зависимости от времени обработки партии (по горизонтальной оси)

Разброс времен обработки партии в зависимости от порядка запуска деталей в обработку является существенным и должен учитываться при поиске решения. В таблице 2 показаны несколько из оптимальных решений.

Таблица 2 – Часть оптимальных решений предложенного примера

№ решения	Порядок запуска деталей в решении							Время обработки партии
	1	2	3	4	5	6	7	
1	2	5	3	6	7	4	1	40
2	2	5	6	3	4	7	1	40
3	3	2	5	4	6	7	1	40
4	3	2	5	7	1	6	4	40

Отметим, что ситуации двух и многих станков существенно различаются в части наличия и возможностей поиска оптимальных решений в случае равенства времен начала и окончания обработки деталей.

2. Результаты и их обсуждение

Правила и алгоритм С. Джонсона неэффективны или неприменимы в ситуациях большого количества станков либо наличия большого количества деталей с равными временами ожидания начала и окончания обработки. Для подобных ситуаций требуется разработать иные подходы поиска оптимальных последовательностей запуска деталей в обработку.

Библиографические ссылки

1. S.M. Johnson. Optimal two- and three-stage production schedules with setup times included // P-402 (RAND Corporation Paper series). Santa Monica: RAND Corp., May 1953. 12 p. URL: https://www.rand.org/content/dam/rand/pubs/papers/2008/P402.pdf
2. Яр-Мухамедов И. Г. Редукция задачи о многих станках // Инновации в технологиях и образовании: сборник статей участников XIII Международной научно-практической конференции, Белово, 26 марта 2020 года. Белово: Кузбасский государственный технический университет имени Т.Ф. Горбачева, 2020. С. 85–87.

ОПТИМИЗАЦИЯ И НАДЕЖНОСТЬ СИСТЕМ

UDC 519.853.3

STRONG DUAL PROBLEMS IN LINEAR COPOSITIVE OPTIMIZATION AND THEIR PROPERTIES

O.I. Kostyukova [a, b], T.V. Tchemisova [c]

[a]*Institute of Mathematics, National Academy of Sciences of Belarus, Surganov str. 11, 220072, Minsk, Belarus,* kostyukova@im.bas-net.by
[b]*Belarusian state University, Independence Ave., 4, 220030 Minsk, Belarus*
[c]*Mathematical Department, University of Aveiro, Campus Universitario Santiago, 3810-193, Aveiro, Portugal,* tatiana@ua.pt

The presentation is devoted to formulation of new strong dual problems for linear copositive optimization and analysis of their properties. Based on the recently introduced concept of the set of immobile indices, here we deduce an extended dual problems that satisfy the strong duality relations and do not require any additional regularity assumptions such as constraint qualifications.

Keywords: Duality Theory; Semidefinite programming; Copositive programming; strong duality.

Introduction

Duality theory is a rich and powerful area of Convex Optimization which is central to understanding sensitivity analysis and infeasibility issues as well as to development of numerical methods.

One says that a pair of dual problems (P) and (D) satisfies *strong duality relation/property* if under the assunption that prime problem (P) is feasible and $val(P) > -\infty$, dual problem (D) has an optimal solution and the duality gap is zero: $val(P) - val(D) = 0$. Here and in what follows, for an optimization problem (P), $val(P)$ denotes the optimal value of its cost function.

Strong duality property is important both from theoretical and practical point of view since the majority of numerical methods (e.g. the Interior Point Method) are based on the assumption that the strong duality relation is satisfied. Violation of this property leads to great numerical difficulties.Hence, it is important to get strong dual formulations that do not need any regularity conditions.

1. Conic optimization problems

A linear conic optimization problems has a (general) form

$$\text{CP:} \qquad \min_x c^\top x \quad \text{s.t. } A(x) \in K,$$

where the decision variable is the $n-$ vector $x = (x_1,...,x_n)^\top$; $\mathcal{A}(x) := \sum_{j=1}^{n} A_j x_j + A_0$ with given matrices $A_j \in S^p$, $j = 0,1,\ldots,n$; S^p is the space of $p \times p$ symmetric matrices, $K \subset S^p$ is a cone. For the conic problem (CP), the classical Lagrangian dual problem has the form

$$\text{LD:} \qquad \max_U -U \bullet A_0, \; s.t. \; U \bullet A_j = c_j, \;\; j = 1,2,\ldots,n, \;\; U \in K^*,$$

where $K^* := \{A \in S^p : A \bullet D \geq 0 \;\; \forall D \in K\}$ is the dual cone of cone K , $A \bullet D = \mathrm{trac}(AD)$.

The most important classes of Conic Optimization are Semidefinite Programming and Copositive Programming problems.

If in the conic problem (CP), we set $K := S_+^p$, we get a problem of Semidefinite Programming (SDP)

$$\text{SDP:} \qquad \min_x c^\top x \quad \text{s.t. } A(x) \in S_+^p.$$

Here $p > 1$, $\mathcal{S}_+^p$ is the cone of symmetric positive semidefinite $p \times p$ matrices.

SDP is a natural generalization of Linear Programming. SDP models have many theoretical and practical applications (see [1, 2]). SDP problems are rather well studied and there are several solvers.

Let $\mathcal{COP}^p \subset \mathcal{S}^p$ be the cone of copositive matrices defined as

$$\mathcal{COP}^p := \{D \in S^p : t^\top Dt \geq 0 \; \forall t \in \mathbb{R}_+^p\}.$$

If in the conic problem (CP) set $K := \mathcal{COP}^p$, then we get a linear Copositive Programming (CoP) problem

$$\text{COP:} \qquad \min_x c^\top x \quad \text{s.t. } A(x) \in \mathcal{COP}^p.$$

CoP is a relatively new field of conic optimization which is actively developed in recent years. CoP problems have important applications (see [3]), including $\mathcal{NP}$ -hard problems.

CoP problems can be considered as a generalization of SDP ones. Unfortunately, CoP problems are less studied than SDP problems. There are many open theoretical problems (see [4]). There exists no methods, nor efficient software for solving CoP problems as well. The reason for this is that the cone $\mathcal{COP}^p$ and its dual, the cone of *completely positive matrices:*

$$\mathcal{CP}^p := \mathrm{conv}\{tt^\top : t \in \mathbb{R}^p_+\}, \tag{1}$$

are more complicated than the cone S^p_+. In fact, the cone S^p_+ is self-dual ($(S^p_+)^* = S^p_+$) and homogeneous (hence symmetric), facially exposed, and nice; but the copositive cone $\mathcal{COP}^p$ *does not* possess *all* mentioned here properties of S^p_+.

One of important open problems for CoP is to formulate dual problems that satisfy strong duality property. The aim of this presentation is to consider and analyze such formulations.

2. Strong dual formulations for SDP and cop problems

In paper [5] for the semidefinite programming problem (SDP), it was obtained the Extended Dual Problem

$$\max -(U + W_{m_0}) \bullet A_0,$$

$$\text{s.t.} \quad (U_m + W_{m-1}) \bullet A_j = 0,\ j = 0,1,...,n,\ m = 1,...,m_0, \tag{2}$$

$$\text{ED-R:} \qquad (U + W_{m_0}) \bullet A_j = c_j,\ j = 1,2,...,n; \quad W_0 = \mathbb{O}_p, \tag{3}$$

$$U \in S^p_+,\ \begin{pmatrix} U_m & W_m \\ W_m^\top & I \end{pmatrix} \in S^{2p}_+,\ m = 1,...,m_0,$$

where $m_0 \geq 0$ is a finite integer, $\mathbb{O}_p$ is the $p \times p$ null matrix

This dual is stated completely in terms of the data of the original program (SDP), and the pair of dual problems (SDP) and (ED-R) satisfies strong duality relation without any additional assumptions.

The aim of our study is to formulate for CoP dual problems which have a form that is similar to the problem (ED-R) and satisfy the strong duality property.

For linear CoP problem (COP), the corresponding Lagrangian dual problem has the form

LDP: $\max_{U} -U \bullet A_0,\ \text{s.t. } U \bullet A_j = c_j,\ j = 1,2,\ldots,n, U \in \mathcal{CP}^p,$

where $\mathcal{CP}^p$ is the dual cone to the cone $\mathcal{COP}^p$ defined in (1).

It is known that under the Slater condition the duality gap for dual pair (COP) and (LDP) is zero. But without the Slater condition, the duality gap can be positive.

Based on an approach proposed in [6, 7], we formulated in [8] a new *extended* dual problem for (COP) in the form

EDP: $\max -(U + W_{m_0}) \bullet A_0,$ s.t. (2), (3) and

$$U \in \mathcal{CP}^p, \begin{pmatrix} U_m & W_m \\ W_m{}^\top & D_m \end{pmatrix} \in \mathcal{CP}^{2p}, m = 1,...,m_0, \tag{4}$$

with the dual variables $U_m \in S^p, W_m \in \mathbb{R}^{p\times p}, D_m \in S^p, m = 1,...,m_0; U \in \mathcal{CP}^p.$

Theorem 1 in [8] justifies that pair of problems (COP) and (EDP) satisfies strong duality relation: if $val(\text{COP}) > -\infty$, then there exists a finite $m_0 \geq 0$ such that the dual problem (EDP) has an optimal solution and $val(\text{COP}) = val(\text{EDP})$.

If we compare the strong duals (ED-R) for problem SDP and (EDP) for CoP problem, we can see that they have a similar structure. The only natural expected difference is that for SDP formulations we use the cone S^p_+ (which is self-dual), and for the CoP formulations use the cone $\mathcal{COP}^p$ (for the primal problem) and the cone $\mathcal{CP}^p$ (which is dual to $\mathcal{COP}^p$) for its dual.

In paper [9], we obtined an estimate for integer m_0 : $0 \leq m_0 \leq p^* = \min\{2n, p(p+1)/2\}$.

For the problem (COP), let us formulate other dual problems and compare their properties. We start with a dual problem that was proposed and justified in our recently submitted paper.

Given a finite integer $m_0 \geq 0$, let us consider the following problem:

DP: $\max -(U + W_{m_0}) \bullet A_0,$ s.t. (2), (3) and

$$U_m \in \mathcal{CP}^p, W_m \in (\mathcal{F}(U_m))^* \ \forall m = 1,...,m_0, \tag{5}$$

where $\mathcal{F}(U) := \{D \in \mathcal{COP}^p : D \bullet U = 0\}$ is the exposed face of $\mathcal{COP}^p$ generated by $U \in \mathcal{CP}^p$.

Theorem 1. *Let the problem* (COP) *be consistent and* $val(\text{COP}) > -\infty$.

Then there exists m_0, $0 \le m_0 \le p^*$ *such that for the pair of problems* (COP) *and* (DP), *the strong duality relations hold true.*

Moreover, one can show that in general the set of feasible solutions of the dual problem (DP) is bigger than the set of feasible solutions of the problem (EDP).

Some other strong dual for Conic Optimization problems was considered in [10]. Theorem 2 from [10], applied to the problem (COP), is as follows.

Theorem 2. *For all large enough integer* m_0, *problem*

$$\text{FDP:} \quad \begin{array}{l} \max(-Y_{m_0+1} \bullet A_0), \quad \text{s.t. } Y_m \bullet A_j = 0,\ j = 0,1,...,n,\ m = 1,...,m_0; \\ \qquad\qquad\qquad\qquad Y_{m_0+1} \bullet A_j = c_j,\ j = 1,2,...,n; \end{array}$$

$$(Y_1, Y_2, ..., Y_{m_0+1}) \in \mathrm{F}R_{m_0+1}(\mathcal{COP}^p) \qquad (6)$$

*is a strong dual for problem (*COP*). Here for integer* $k \ge 1$, $\mathrm{F}R_k(\mathcal{K})$ *denotes a facial reduction cone of order* k of a cone $\mathcal{K}$:

$$\mathrm{F}R_k(\mathcal{K}) := \{(Y_1, Y_2, ..., Y_k) : Y_1 \in \mathcal{K}^*, Y_m \in (\mathcal{K} \cap Y_1^{\perp} \cap ... \cap Y_{m-1}^{\perp})^*, m = 2,...,k\}.$$

Thus, the variables of the problem (FDP) (the dual variables) belong to the facial reduction cone of order $m_0 + 1$ of the cone $\mathcal{COP}^p$. It was shown in [10] that for any $k \ge 1$, the cone $\mathrm{F}R_k(\mathcal{COP}^p)$ is convex and, for any $k > 1$, it is not closed.

The following lemma shows that the set of feasible solutions of the problem (FDP) is wider than the set of feasible solutions of the problem (DP).

Lemma 1. *Let* $(W_0, U_m, W_m, m = 1,...,m_0, U)$ *be a feasible solution of the problem (*DP*). Then* $(Y_1 = U_1, Y_m = U_m + W_{m-1}, m = 2,...,m_0, Y_{m_0+1} = U + W_{m_0})$ *is a feasible solution of the problem* (FDP).

Thus we have considered several dual problems for the copositive problem (COP) that satisfy strong duality relation without any additional assumptions. Having compared these dual problems, we can state the following.

1) The problems (EDP), (DP), and (FDP) differ from each other in constraints (4), (5), and (6).

2) The problem (EDP) can be considered as a completely positive problem. The problems (DP) and (FDP) are conic problems whose variables belong to the cones $\mathrm{F}R_2(\mathcal{COP}^p)$ and $\mathrm{F}R_{m_0+1}(\mathcal{COP}^p)$, respectively.

3) The dual problems (EDP) and (DP), contain m_0 separate explicite conditions (4) and (5), respectively, for each $m = 1,...,m_0$. In the problem (FDP),

instead of these m_0 constraints, there is a unique, but more complex constraint (6) in a recursive form (this constraint can be considered as a kind of "aggregation" of the mentioned above "simple" constraints in the problem (EDP)).

4) The facial reduction cone $\mathrm{F}R_{m_0+1}(\mathcal{COP}^p)$ used in the problem (FDP) is not explicitly described. The dimension of this cone is large, which greatly complicates the solution of this problem.

5) Each feasible solution of the problem (EDP) generates a feasible solution of the problem (DP), and each feasible solution of the latter problem generates a feasible solution of the problem (FDP).

Conclusions

The main contribution of the presentation consists in considering some new dual problems for the copositive problem and comparing their properties.The results provide templates for creating other strong dual formulations for linear/convex copositive problems. These formulations can be used for a variety of purposes, both theoretical and practical.

References

1. Vandenberghe L., Boyd S. Semidefinite Programming // SIAM Review. 1996. № 38. P. 49–95.
2. Wolkowicz H., Saigal R., Vandenberghe L. Handbook of Semidefinite Programming: Theory, Algorithms, and Applications. Springer Science & Business Media,.2000, 654 p.
3. Bomze I.M. Copositive optimization - Recent developments and applications // *EJOR*. 2012. № 216(3). P. 509–520.
4. Berman A., Dür M., Shaked-Monderer N. Open problems in the theory of completely positive and copositive matrices // Electronic Journal of Linear Algebra. 2015. № 29. P. 46–58.
5. Ramana M.V., Tuncel L., Wolkowicz H. Strong duality for Semidefinite Programming // SIAM J. Optimization. 1997. № 7(3). P. 641–662.
6. Kostyukova O.I., Tchemisova T.V. On a constructive approach to optimality conditions for convex SIP problems with polyhedral index sets // Optimization. 2014.№ 63(1). P. 67–91.
7. Kostyukova O.I., Tchemisova T.V. Optimality conditions for convex Semi-Infinite Programming problems with finitely representable compact index sets // J. Optim.

Theory Appl. 2017. № 175(1). P. 76–103.
8. Kostyukova O., Tchemisova T. On strong duality in linear copositive programming // J Glob Optim. 2022. № 83. P. 457–480.
9. Kostyukova O., Tchemisova T. An exact explicit dual for the linear copositive programming problem. Optimization Letters. 2022. DOI:10.1007/s11590-022-01870-0.
10. M. Liu, G. Pataki G. Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming // Math. Program. 2018. № 167. P. 435–480.

УДК 519.71

РАЗРАБОТКА БАЗЫ ДАННЫХ ПО ДЕСКРИПТОРНЫМ СИСТЕМАМ

И.К. Асмыкович, Д.Е. Сидорчик, А.А. Королев

Белорусский государственный технологический университет, ул. Свердлова 13-а, 220006, г. Минск, Беларусь, asmik@tut.by

Доклад посвящен результатам многолетней работы по сбору публикаций по специальному классу математических моделей систем управления – дескрипторным системам, которые ряд авторов называют дифференциально-алгебраическими, либо сингулярными, либо неразрешенными относительно производной, либо гибридными. Библиографический указатель составлен на основе анализа публикаций в многочисленных журналах по математической теории управления, реферативных журналов, специализированных информационных ресурсов Интернета, списков цитированных работ. Для удобства работы с указателем составлена база данных, которая позволяет быстро находить работы конкретного автора, распределение публикаций по годам, работы по специальным задачам для таких систем.

Ключевые слова: Библиографический указатель; теория управления; дескрипторные системы; база данных.

DEVELOPMENT OF A DATABASE FOR DESCRIPTOR SYSTEMS

I.K. Asmykovich, D.E. Sidorchik, A.A. Korolev

Belarusian State Technological University, st. Sverdlova 13-a, 220006, Minsk, Belarus, asmik@tut.by

The report is devoted to the results of many years of work on collecting publications on a special class of mathematical models of control systems - descriptor systems, which a number of authors call differential-algebraic, or singular, or unresolved with respect to the derivative, or hybrid. The bibliographic index is compiled on the basis of an analysis of publications in numerous journals on mathematical control theory, abstract journals, specialized Internet information resources, and lists of cited works. For the convenience of working with the index, a database has been compiled that allows you to quickly find the works of a particular author, the distribution of publications by year, and works on special problems for such systems.

Keywords: Bibliographic index; control theory; descriptor systems; database.

Введение

В качественной теории управления движением, которая активно развивалась в XX веке, основной математической моделью часто была линейная система обыкновенных дифференциальных уравнений или урав-

нений в частных производных, или нелинейная модель для которой обычно рассматривалось линейное приближение. Для таких объектов подробно рассмотрены и проанализированы основные задачи качественной теории управления, получены критерии их разрешимости, разработаны алгоритмы синтеза необходимых регуляторов. Основные публикации до 80-х годов отражены в библиографических указателях [1,2], основные результаты во многочисленных обзорах, например, [3]. Но в XXI веке было выяснено, что даже для таких моделей решены далеко не все задачи [4].

Классическим математической моделью в качественной теории управления для обыкновенных линейных систем является система вида

$$\begin{aligned} &\dot{x}(t) = Ax(t) + Bu(t), \\ &x(0) = x_0, \end{aligned} \tag{1}$$

$$y(t) = Cx(t) \tag{2}$$

где $x(t)$ – n-вектор состояния, $u(t)$ – r - вектор управляющих воздействий, у – m-вектор выхода или наблюдаемых координат, A, B, C – постоянные матрицы соответствующих размеров.

При дальнейшем изучении реальных динамических систем управления было выяснено, что представление (1), (2) далеко не всегда корректно описывают объект управления. При составлении математических моделей физических процессов и систем автоматического регулирования необходимо учитывать, как дифференциальные, так и алгебраические связи, как непрерывные, так и дискретные взаимодействия. Адекватной математической моделью таких процессов являются линейные системы дифференциальных уравнений неразрешенные относительно производной. На необходимость изучения таких систем обращал внимание академик Лузин Н.Н. [5]. Такие системы находят широкое распространение в самых разнообразных областях современной науки и техники: в автоматике и телемеханике, радиологии, биологии и медицине, при моделировании технологических процессов в плазме и лазерах, ряде экономических моделей и т.д. [7-16]. Их называют дескрипторными [11, 13, 14,15]. а также дифференциально-алгебраическими [12]. или алгебро-дифференциальными [9, 10]. либо сингулярными [8, 16], либо неразрешенными относительно производной [7]. Математически такие системы записываются в виде

$$\begin{aligned} &S\dot{x}(t) = Ax(t) + Bu(t), \\ &Sx(0) = Sx_0, \quad \det S = 0, \end{aligned} \tag{3}$$

В таких системах имеются существенные сложности в вопросах су-

ществования и единственности решения, но при условии регулярности

$$\det[\lambda S - A] \neq 0 \qquad (4)$$

они снимаются.

Системы в виде (3) широко используются при моделировании процессов в электрических цепях, технологических процессов переноса (материала и тепла), задачах демографии. Это происходит тогда, когда наряду с дифференциальными связями встречаются и алгебраические (функциональные) зависимости, например, условия материального или финансового баланса.

Отметим, что при составлении указателей прежними способами возникают трудности при их использовании. Так указатель [5] содержит 6016 публикаций, но найти работы по конкретным задачам теории непрерывных дробей, или посмотреть распределение работ по годам весьма затруднительно.

Понятно, что при этом возникает проблема дублирования работ и исследований, когда одни и те же задачи с небольшими изменениями изучаются в большом количестве статей или тезисов и материалов конференций. Поэтому возникла идея использовать современные возможности информационных технологий и разработать базу данных, которая позволит быстро находить публикации конкретного автора, даже если статья в соавторстве, выяснять число и годы публикаций по конкретным задачам для дескрипторных систем.

1. Методология исследования / теоретические основы

Данная база данных создана на основе библиографического указателя [15] по теме дескрипторных систем управления. Указатель составлялся в течении 35 лет путем анализа реферативных журналов по данной тематике, просмотра материалов научных конференций по качественной теории управления, анализа списков цитирования статей по дескрипторных системах. База данных содержит данные о литературных источниках, журналах, материалах конференций, монографиях и сборниках научных трудов в которых были опубликованы статьи по указанной тематике, а также полный список авторов. Реализована в СУБД Microsoft Access.

Данные расположены в основных трёх таблицах: «Журналы», «Авторы», «Литература». Для осуществления возможности выборки данных между ними использовались связи «один-ко-многим» (реализуется тогда, когда объекту А может принадлежать или же соответствовать несколько

объектов Б, но объекту Б может соответствовать только один объект А) и «многие-ко-многим» (реализуется в том случае, когда нескольким объектам из таблицы А может соответствовать несколько объектов из таблицы Б, и в тоже время нескольким объектам из таблицы Б соответствует несколько объектов из таблицы А). «Журналы» и «Литература» связаны первым типом связи, «Литература» и «Авторы» - вторым. Для реализации связи «многие-ко-многим» была создана промежуточная таблица, содержащая в себе ключи связанных таблиц, которые являются значениями, уникально идентифицирующие каждую запись в них.

Для удобства отображения информации для пользователя было создано 3 запроса на выборку данных, которые в зависимости от введённой информации пользователем выводят соответствующие записи в объединённой таблице. Реализовано 3 варианта отбора записей:

- выборка данных на основе информации о публикации;
- выборка данных на основе информации об авторе;
- выборка данных для отображения публикаций в журнале.

При выборке данных на основе информации о публикации или об авторе, вводится непосредственно полное наименование или фамилия, либо отдельная фраза или буква в первом диалоговом окне. Также потом можно ввести конкретный год, если пользователя интересуют публикации за какой-то конкретный промежуток времени. В выборке данных по журналу просто достаточно ввести наименование журнала или ключевое слово, по которым будут выведены публикации в подходящих под условия запроса журналах.

Для удобства восприятия информации для пользователя в выборки по авторам и журналам была добавлена функция-модуль, написанная на языке Visual Basic, осуществляющая подсчёт выведенных записей и присвоению каждой из них в выборке своего порядкового номера. Модуль – это объект Access, в котором хранится коллекция процедур. Внутри данного модуля написана пользовательская функция Numeration, которая принимает в себя один аргумент var, по которому и осуществляется нумерация записей. Функция, в свою очередь, состоит из ряда операторов, которые выполняют некоторое действие, и тоже могут получать аргументы. Однако, в отличие от подпроцедур в Access, функции возвращают значение.

В дальнейшем данная база может быть улучшена, в частности, может быть добавлен графический интерфейс для более понятной и удобной работы с данными для пользователя, добавлены возможности одновременного поиска по фамилии и журналу, году и автору, году и журналу

2. Результаты и их обсуждение

Таким образом если надо выяснить современное состояние исследований по стабилизации или модальному управлению, управляемости или наблюдаемости, расщепимости или реконструкции дескрипторных систем, то это можно будет сделать достаточно быстро. Такая база данных позволит при переходе на изучение специальных классов дескрипторных систем, в частности, дескрипторных систем с запаздыванием, систем над коммутативными кольцами, нелинейных систем различных классов, систем с многомерным временем находить мгновенно список публикаций. Можно получить список научных журналов и конференций, где рассматриваются системы такого типа. Конечно, сейчас в Интернете есть большое количество поисковых систем, но все они весьма широкого профиля, и различают дифференциально – алгебраические, сингулярные, неразрешенные относительно производной, а в данной базе данных они объединены и поэтому собраны почти все работы по дескрипторным системам за 40 лет [14, 15].

Библиографические ссылки

1. Теория управления движением. Часть I. Линейные конечномерные системы. Библиогр. указ. / Сост. Габасов Р. Кириллова Ф.М., Марченко В.М., Асмыкович И.К. Мн.: Ин-т матем. АН БССР, 1983. 132 с.
2. Теория управления движением. Часть II. Нелинейные и бесконечномерные системы: Библиогр. указ. / Сост. Габасов Р. Кириллова Ф.М., Марченко В.М., Асмыкович И.К. Мн.: Ин-т матем. АН БССР, 1983. 87 с.
3. Асмыкович И.К., Габасов Р., Кириллова Ф.М., Марченко В.М. Задачи управления конечномерными системами // Автоматика и телемеханика. 1986. №11. С. 5–29.
4. Поляк Б.Т., Щербаков П.С. Трудные задачи линейной теории управления // Автоматика и телемеханика 2005, № 5. С.7-46.
5. Лузин Н.Н. К изучению матричной теории дифференциальных уравнений // Автоматика и телемеханика. 1940. № 5. С. 4–66.
6. Шмойлов В. И., Коровин Я.С., Войтулевич В. Ю. Непрерывные дроби. Библиографический указатель. Lap Lambert Academic Publishing, Saarbruken, Germany, 2017. 560 с.
7. Ахундов А.А. Обзор некоторых результатов по теории линейных дифференциальных уравнений, неразрешенных относительно производной // Mathematical control theory / Banach Contr. Publ. Warsaw, 1985. Vol. 14. P. 7–16.

8. Dai L. Singular Control Systems // Lecture Notes in Control and information Sciences, Vol.118. Berlin, Springer-Verlag, 1989. 439 p.

9. Чистяков В. Ф., Щеглова А. А. Избранные главы теории алгебро-дифференциальных систем. Новосибирск: Сибирская издательская фирма РАН "Наука", 2003. 320 с.

10. Бояринцев Ю.Е. Линейные и нелинейные алгебро-дифференциальные системы. Новосибирск: Наука. 2000.

11. Feng Yu, Yagoubi M. Robust Control of Linear Descriptor Systems Publisher : Springer Singapore, Mar. 2017.

12. Ilchmann A., Reis T. Surveys in Differential-Algebraic Equations I-IV Differential-Algebraic Equations Forum, Berlin, Heidelberg, Springer, 2013-2017.

13. Белов А. А., Курдюков А. П. Дескрипторные системы и задачи управления. М.: Физматлит, 2015. 272 с.

14. Дескрипторные системы управления: Библиогр. Указ. / АН БССР, Ин-т математики: Сост.: Р. Габасов, Ф. М. Кириллова, И. К. Асмыкович. Мн., 1988. 38 с.

15. Дескрипторные системы управления: Библиографический указатель / сост. И.К Асмыкович. Минск: БГТУ, 2022. 343 с

16. Debeljkovic D. Lj, I.. M. Buzurovic Lyapunov Stability of Linear Continuous Singular Systems: An Overview // International Journal of Information & System Science, (Canada), 2011, V. 7, №. 2-3. P. 247–268.

УДК 517.977

К УПРАВЛЯЕМОСТИ ЛИНЕЙНЫХ НЕСТАЦИОНАРНЫХ ДИСКРЕТНЫХ СИСТЕМ С ИНТЕРВАЛЬНЫМИ НЕОПРЕДЕЛЕННОСТЯМИ

В.В. Горячкин, В.В. Крахотко, Г.П. Размыслович

Белорусский государственный университет, пр.Независимости, 4, 220030, г.Минск, Беларусь, gorvv@bsu.by, krakhotko@bsu.by, razmysl@bsu.by

Исследована задача управляемости линейных нестационарных дискретных динамических систем с интервальными неопределенностями. Для таких систем вводятся понятия скользящего окна управляемости и запаса управляемости. Для однородной системы с интервальной матрицей строится матрица Коши и описываются ее свойства, исходя из понятий интервального анализа. Получены достаточные условия управляемости, записанные в терминах сингулярных чисел матрицы управляемости.

Ключевые слова: нестационарная дискретная система; управляемость систем; интервальный анализ; сингулярные числа матрицы.

ON CONTROLLABILITY PROBLEM OF LINEAR NONSTATIONARY DISCRETE SYSTEMS WITH INTERVAL INDETERMINATIONS

V.V. Goryachkin, V.V. Krakhotko, G.P. Razmyslovich

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus
Corresponding author: gorvv@bsu.by, krakhotko@bsu.by, razmysl@bsu.by

The controllability problem of linear nonstationary discrete dynamical systems with interval indeterminations is investigated. For such systems, the concepts of a sliding window of controllability and a margin of controllability are introduced. Basing on the concepts of interval analysis a Cauchy matrix for a homogeneous system with an interval matrix is constructed and its properties are described. Sufficient controllability conditions written in terms of singular numbers of the controllability matrix are obtained.

Keywords: nonstationary discrete system; controllability; interval analysis; singular matrix numbers.

Введение

Как известно, что задача управляемости динамических систем предшествует задаче оптимального управления. Поэтому задачи управляемости систем актуальны и важны для динамических систем и для систем управления с интервальными неопределенностями, которые описывают различные прикладные задачи в экономике, технике и т.д.

Реальные системы управления функционируют, как правило, в условиях неопределенности, обусловленной разнообразными причинами (неполные начальные данные, наличие неизвестных и неточно заданных возмущений, ошибки в каналах связи и т.п.). Поэтому в реальных задачах существуют неопределенности в описании объекта. В тоже время реальная система управления должна быть спроектирована и реализована так, чтобы она была работоспособна при наличии этих неопределенностей. В докладе описывается подход к задачам управляемости континуума дискретных динамических систем в условиях существенно ограниченной исходной информации.

1. Вспомогательные результаты

Пусть заданы нестационарные матрицы $A(t) \in R^{n\times n}$ и $B(t) \in R^{n\times r}$ ($t \in Z_+$) с интервальными неопределенностями соответственно $A(t) \in [A(t)]$, $B(t) \in [B(t)]$.

Введем в рассмотрение интервальные матрицы $[P^1(i,j)]$, $[D^1(i,j)]$:

$$P^1(i,j) = \begin{cases} A(i-1)A(i-2)\ldots A(j) & \text{при} \quad i > j \geq 0 \\ E & \text{при} \quad i = j \end{cases},$$

$$P_0^1(i,j) = \begin{cases} A_0(i-1)A_0(i-2)\ldots A_0(j) & \text{при} \quad i > j \geq 0 \\ E & \text{при} \quad i = j \end{cases},$$

$D^1(i,j) = P^1(i,j)B(j-1)$ ($i > j \geq 0$), $D_0^1(i,j) = P_0^1(i,j)B_0(j-1)$ ($i > j \geq 0$),

причем:

a) при $A_i = A(i)$ определяем центральную (номинальную) матрицу интервальной матрицы $[P^1(i,j)]$ как $mid([P^1(i,j)]) = P_0^1(i,j)$, а радиус как

$$rad([P(i,j)]) = \Delta P(i,j) = \prod_{k=1}^{i-j} abs([A(i-k)]) - |P_0(i,j)| \geq 0 .$$

b) при $A_i = A(i)$, $Bj = B(j-1)$ положим в качестве центра интервальной матрицы $[D^1(i,j)]$ - $mid([D^1(i,j)]) = D_0^1(i,j) = P_0^1(i,j)B_0(j-1)$, а радиуса -

$$rad([D^1(i,j)]) = \Delta D^1(i,j) = \left(\prod_{k=1}^{i-j} abs([A(i-k)]) \right) abs([B(j-1)]) - \left| P_0^1(i,j)B_0(j-1) \right| \geq 0.$$

Ясно, что для любых матриц $A(t) \in [A(t)]$, $B(t) \in [B(t)]$. имеем $P^1(i,j) \in [P^1(i,j)]$, $D^1(i,j) \in [D^1(i,j)]$.

Так как в интервальном анализе операция умножения не обладает свойством ассоциативности, т.е. ([A][B])[C]≠[A]([B][C]), то определим произведение ([A][B])[C] – как левое, а [A]([B][C]) – как правое произведение.

Введем в рассмотрение матрицы с правым и левым матричным произведением. Положим интервальные матрицы с правым произведением:

$$[P^2(i,j)] = \begin{cases} [A(i-1)]([A(i-2)](\cdots([A(j+1)][A(j)]))\cdots) & \text{при } i > j \geq 0, \\ E, \text{ при } i = j \end{cases},$$

$$[D^2(i,j)] = \begin{cases} [A(i-1)]([A(i-2)](\cdots(A(j+1)([A(j)][B(j-1)]))\cdots) & \text{при } i > j \geq 0, \\ E, \text{ при } i = j \end{cases}$$

Тогда, легко видеть, справедливы включения

$$\left\{A(i)P^1(i,j) = P^1(i+1,j) \middle| A(i) \in [A(i)], P^1(i,j) \in [P^2(i,j)]\right\} \subseteq$$

$$\subseteq \left\{P^2(i+1,j) \middle| P^2(i+1,j) \in [P^2(i+1,j)] = [A(i)][P^2(i,j)]\right\},$$

$$\left\{A(i)D^1(i,j) = D^1(i+1,j) \middle| A(i) \in [A(i)], D^1(i,j) \in [D^2(i,j)]\right\} \subseteq$$

$$\subseteq \left\{D^2(i+1,j) \middle| D^2(i+1,j) \in [D^2(i+1,j)] = [A(i)][D^2(i,j)]\right\}.$$

Введем по аналогии матрицы с левым произведением. Положим

$$[P^3(i,j)] = \begin{cases} (\cdots(([A(i-1)])[A(i-2)])[A(i-3)])\cdots)[A(j+1)][A(j)] & \text{при } i > j \geq 0, \\ E, \text{ при } i = j \end{cases},$$

$$[D^3(i,j)] = \begin{cases} (\cdots(([A(i-1)])[A(i-2)])[A(i-3)])\cdots)[A(j+1)])[A(j)])[B(j-1] & \text{при } i > j \geq 0, \\ E \text{ при } i = j \end{cases}$$

Легко видеть, что имеет место равенство $[D^3(i,j)] = [P^3(i,j)][B(j-1)]$. И тогда, очевидно, имеют место включения.

$$\left\{A(i)P^1(i,j) = P^1(i+1,j) \middle| A(i) \in [A(i)], P^1(i,j) \in [P^3(i,j)]\right\} \subseteq$$

$$\subseteq \left\{P^3(i,j-1) \middle| P^3(i,j-1) \in [P^3(i,j-1)] = [P^3(i,j)][A(j-1)]\right\},$$

$$\left\{P^1(i,j)B(j-1) = D^1(i,j) \middle| B(j-1) \in [B(j-1)], P^1(i,j) \in [P^3(i,j)]\right\} \subseteq$$

$$\subseteq \left\{D^3(i,j) \middle| D^3(i,j) \in [D^3(i,j)] = [P^3(i,j)][B(j-1)]\right\}.$$

Результат пересечения указанных матриц - интервальные матрицы:

$[P(i,j)]=[P^1(i,j)]\cap[P^2(i,j)]\cap[P^3(i,j)]$,

$[D(i,j)]=[D^1(i,j)]\cap[D^2(i,j)]\cap[D^3(i,j)]$,.

содержащие соответственно точечные матрицы $P^1(i,j), D^1(i,j)$, когда матрицы с интервальными неопределенностями $A(t)$, $B(t)$ независимо друг от друга пробегают любые значения из заданных интервалов $A(t)\in[A(t)]$, $B(t)\in[B(t)]$.

Воспользуемся, полученными результатами для исследования задач управляемости.

2. Управляемость по состоянию

Рассмотрим нестационарную дискретную систему

$$x(t+1)=A(t)x(t)+B(t)u(t), t\in Z_+ \qquad (1)$$

с интервальными неопределенностями $A(t)\in[A(t)]$, $B(t)\in[B(t)]$. Здесь $x(t)\in R^n-$ вектор состояния; $u(t)\in R^r-$ управляющее воздействие; $A(t)$, $B(t)$ и $[A(t)],[B(t)]$ соответственно точечные и интервальные вещественные матрицы согласованной размерности.

Пусть η,τ целые числа, причем $\tau=\eta+\rho,\rho\geq 1$ и $(\eta,\tau)=\{\eta,\eta+1,\ldots,\tau-1\}=\{\eta,\eta+1,\ldots,\eta+\rho-1\}$ – подмножество (называемое отрезком) множества Z_+, $\rho-$ диаметр этого подмножества (ширина отрезка).

Определение 1. *Система* (1) *управляема по состоянию на множестве* (η,τ) [2], *если для любых точек* v,w *из* R^n *найдется такое управление* $u(\eta,\tau)=\{u_\eta,u_{\eta+1},\ldots,u_{\tau-1}\}$, *что* $x(t)=x(t,\eta,u(\eta,\tau))$ *в момент* $t=\tau$ *удовлетворяет условию* $x(\tau)=w$ *(здесь* $x(t)=x(t,\eta,u(\eta,\tau))$ *- решение системы* (1)*, порожденное начальным значением* $x(\eta)=v$ *и входным воздействием* $u(\eta,\tau)$*).*

Определение 2. *Скользящим окном управляемости* (η,τ) *(далее просто окном управляемости) назовем множество значений независимой переменной* $\eta\leq t\leq\tau$, *на которых система* (1) *с центральными матрицами* $A_0(t)$ и $B_0(t)$ (*так называемая центральная система) вполне управляема, стартуя с начального состояния* $x(\eta)$. *Другими словами, стартуя с произвольного* $x(\eta)$, *траектория решения центральной (номинальной)*

системы в момент времени τ *попадет в любое наперед заданное состояние.*

Определение 3. *Интервальная система управляема по состоянию в окне управляемости тогда и только тогда, когда управляемо по состоянию каждое уравнение системы в этом окне.*

Пусть (η,τ) некоторое окно управляемости. Очевидно, матрица $P^1(i,j)$ является матрицей Коши [2] однородного уравнения, полученного из (1) при $u(t)\equiv 0, t\geq 0$. Согласно формуле Коши [2], элемент $x(t)$ в любой момент t в окне (η,τ) определяется соотношением

$$x(t)=P^1(t,\eta)v+\sum_{j=\eta}^{t-1}P^1(t,j+1)B(j)u(j)=P^1(t,\eta)v+\sum_{j=\eta}^{t-1}D^1(t,j+1)u(j),\ x(\eta)=v.$$

В окне (η,τ) определим матрицу $Q^1(\tau,\eta)=$ $=(D^1(\tau,\eta+1),D^1(\tau,\eta+2),\ldots,D^1(\tau,\tau))$. Известно, что система (1) управляема на множестве (η,τ) тогда и только тогда, когда $rank(Q^1(\eta,\tau))=n$ [2].

Рассмотрим интервальную матрицу

$$[Q(\tau,\eta)]=([D(\tau,\eta+1)],[D(\tau,\eta+2)],\ldots,[D(\tau,\tau)]).$$

Определение 4. *Интервальную матрицу назовем матрицей полного ранга* [3]*, если она содержит точечные матрицы только полного ранга.*

Из определения 4 следует, если в окне (η,τ) $[Q(\tau,\eta)]$ является матрицей полного ранга, то множество всех матриц управляемости интервальной системы в этом окне так же состоит только из матриц полного ранга.

Для установления достаточных условий управляемости воспользуемся признаками полного ранга интервальных матриц, приведенными в работе [3].

Отсюда справедлива

Теорема 1. *Интервальная система (1) управляема по состоянию в окне* (η,τ)*, если выполнено хотя бы одно из следующих условий:*

a) $\sigma_{\max}(rad([Q(\tau,\eta)]))<\sigma_{\min}(mid([Q(\tau,\eta)]))$; (2)

b) система неравенств $\left|z'mid([Q(\tau,\eta)])\right|\leq\left|z'\right|rad([Q(\tau,\eta)]), z\in R^n$, (3)

имеет единственное нулевое решение относительно вектор-столбца z.

Здесь $\sigma_{\max}(.)$ *и* $\sigma_{\min}(.)$ *- наименьшее и наибольшее из сингулярных чисел матрицы* [4].

Доказательство. Согласно теореме и следствию [3] выполнение, по крайней мере, одного из условий (2), (3) влечет полноранговость ин-

тервальной матрицы $[Q(\tau,\eta)]$. Значит, интервальная система (1) управляема.

Имеет место

Теорема 2. *Пусть центральная система* ($A_0(t) = mid([A(t)])$, $B_0(t) = mid([B(t)])$)

$$x(t+1) = A_0(t)x(t) + B_0(t)u(t) \qquad (4)$$

управляема по состоянию в окне (η,τ) *(то есть матрица управляемости центральной системы (4)* $Q_0^1(\tau,\eta) = (D_0^1(\tau,\eta+1), D_0^1(\tau,\eta+2),\ldots, D_0^1(\tau,\tau)) = mid([Q^1(\tau,\eta)])$ *полного ранга), тогда каждое уравнение интервальной системы управляемо по состоянию, если*

$$\rho(rad([Q^1(\tau,\eta)])\left|(mid([Q^1(\tau,\eta)]))^+\right| < 1,$$

где $\rho(\cdot)$ - *спектральный радиус;* $(mid([Q^1(\tau,\eta,)]))^+$ - *псевдообратная* [4] для *точечной матрицы* $mid([Q^1(\tau,\eta)])$.

Доказательство теоремы следует из доказательства достаточного условия полноранговости интервальной матрицы $[Q^1(\tau,\eta)]$ ([3, с. 295-296]).

Замечание 1. Заметим, что матрица управляемости центральной системы (4) $mid([Q^1(\tau,\eta)])$ может не совпадать с серединой интервала $[Q(\tau,\eta)]$.

Для быстрого и грубого оценивания спектрального радиуса, который используется в теореме 2, можем использовать его оценку сверху какой-нибудь матричной нормой.

Замечание 2. В работе [3] приведены примеры интервальных матриц, которые по одним признакам имеют полный ранг, по другим нельзя сделать определенный вывод. Примеры также показывают, что для интервальных матриц традиционные способы рассуждений линейной алгебры и матричного анализа и сформировавшаяся интуиция на их основе могут не работать. Например, в обычном не интервальном случае полноранговая матрица просто по определению имеет квадратную неособенную матрицу, порядок которой равен рангу матрицы. В интервальном случае это не так. Поэтому для определения полноранговости интервальной матрицы рекомендуем воспользоваться предложенными признаками.

Пример 1. Интервальная матрица $\begin{pmatrix} 1 & -1 & [-1,1] \\ [0,1] & [0,1] & 1 \end{pmatrix}$ – полного ранга (ранг матрицы равен 2), в тоже время она не содержит неособенных

интервальных 2×2 – подматриц, что нетрудно обнаружить полным перебором всех таких подматриц.

Замечание 3. Увеличение продолжительности воздействия управляющего воздействия (расширение окна (η,τ)), вообще говоря, приводит к возрастанию вероятности наличия свойства управляемости уравнения (1). Как показывают примеры это, вообще говоря, имеет место и в случае управляемости интервальной системы (1).

3. Запас управляемости интервальной системы

На практике при проектировании реальных систем управления, математические модели, которых задаются дискретным уравнением вида (1), возникает задача о сохранении управляемости системы, когда ее параметры независимо друг от друга изменяются в некоторых множествах. Естественно, возникает задача оценки диаметров множеств неопределенности параметров (назовем – порогов допустимой разбалансировки). Математически это можем описать следующим образом.

Пусть определено некоторое окно (η,τ), в котором центральная система управляема. Радиусы интервальных матриц $\Delta A(t)$ и $\Delta B(t)$ будем считать порогами допустимой разбалансировки, при которых система сохраняет свойство управляемости.

Запас управляемости интервальной системы определим как максимальное вещественное число $\alpha\geq 0$ (коэффициент пропорциональности при радиусах интервальных матриц), при котором интервальная система управляема с неопределенностями $A(t)\in[A_\alpha(t)]=[A_0(t)-\alpha\Delta A(t),$ $A_0(t)+\alpha\Delta A(t)],\ B(t)\in[B_\alpha(t)]=[B_0(t)-\alpha\Delta B(t),B_0(t)+\alpha\Delta B(t)]$.

Оценку и нахождение запаса управляемости α проводим в четыре этапа по следующей схеме:

1. По матрице управляемости центральной (номинальной) системы находим ее первое окно управляемости (η,τ) и фиксируем его.
2. Если требование полного ранга интервальной матрицы $[Q^1(\tau,\eta)]$ не выполняется в установленном окне при $\alpha=1$, то, скорей всего, были просчеты на стадии проектирования динамической системы при выборе допустимых порогов разбалансировки. Будем искать $\alpha\in[0,1)$ пропорционально уменьшая радиусы $\Delta A(t)$ и $\Delta B(t)$, например, по аналогии с методом половинного деления отрезка пополам. Итерационный процесс уменьшения радиусов матриц, очевидно, всегда сходится, так как при $\alpha=0$ интервальная система вырождается в управляемую центральную систему.

3. Если проектировщика не устраивают полученные в пункте 2 радиусы $\alpha\Delta A(t)$ и $\alpha\Delta B(t)$, то переходим на пункт 1 для поиска следующего окна управляемости. Иначе выход.
4. Если требование управляемости центральной системы выполняется в окне при $\alpha = 1$, то выберем некоторое $\Delta\alpha > 0$, и на первом шаге проверим требование полного ранга интервальной матрицы $[Q^1(\tau,\eta)]$ с коэффициентом $\alpha + \Delta\alpha$. Далее продолжим на каждом шаге увеличивать α на $\Delta\alpha$ до тех пор, пока требование полноранговости интервальной матрицы будет не выполнено. В качестве запаса управляемости берем предыдущее значение α. Если точность оценки не устраивает, то вернемся на предыдущее значение α и продолжим процесс уточнения (по схеме пункта 2), уменьшая $\Delta\alpha$ на 2. Так на некоторой итерации k с $\Delta\alpha/2^k$ будет найдено $\alpha^* > 1$ с наперед заданной точностью, при котором на интервалах неопределенности $[A_{\alpha^*}(t)]$ и $[B_{\alpha^*}(t)]$ интервальная система сохраняет свойство управляемости в исследуемом окне. Если проектировщика устраивают на некотором шаге увеличенные радиусы $\alpha\Delta A(t)$ и $\alpha\Delta B(t)$ (допустимые пороги разбалансировки), то процесс оценки запаса управляемости завершаем.

Если выбранное окно управляемости не устраивает проектировщика или процедура оценки запаса управляемости не привела к желаемому результату, то ищем следующее окно управляемости и оценку запаса управляемости ансамбля проводим по предложенной выше схеме.

Пример 2. Рассмотрим модельный пример, в котором для простоты ограничимся случаем $n = 2, r = 1$. Пусть соответствующие матрицы системы (1) определены на интервалах

$$[A(t)] = \begin{pmatrix} [2\sin\frac{\pi}{4}\mathrm{t},\ \mathrm{t}(\mathrm{t}+1)] & [\mathrm{t},\ \exp(\mathrm{t})] \\ [t-1,\ t+1] & [t\cos\frac{\pi}{3}\mathrm{t},\ \mathrm{t}+0.5\ ] \end{pmatrix}, [B(t)] = \begin{pmatrix} [\sqrt{t}-1,\ t+1] \\ [t^2-1,\ t^2+1] \end{pmatrix}, t \in Z_+.$$

Для уменьшения объема вычислений сначала найдем первое окно, в котором центральная система управляема. Это будет окно (1,2), и проверим требование управляемости ансамбля. В этом окне центральная система управляема, но требование полноранговости интервальной матрицы $[Q^1(2,1)]$ не выполняется. И только в окне (3,4) обнаружили, что центральная система управляема с матрицей управляемости $Q_0^1(4,3)$, то есть

$$rank(Q_0^1(4,3)) = rank\begin{pmatrix} 287.355\ 7 & 3.00 \\ 20.681\ 02 & 16.0 \end{pmatrix} = 2,$$

и соответствующая интервальная 2×2 – матрица $[Q^1(4,3)]$ – полного ранга

$$rank([Q^1(4,3)]) = rank\begin{pmatrix} [\ 32.002331806,\ 625.98150033] & [\ 1.00,\ 5.00\ ] \\ [\ -17.877363946\ ,\ 65.00000000\ ] & [\ 15.0\ ,\ 17.0\ ] \end{pmatrix} = 2.$$

Кстати, на отрезке [0,2] можем вообще отключить управление и включить его вычисление только в окне управляемости (3,4).

В окне (3,4) оценим с заданной точностью верхнюю границу запаса управляемости. Будем последовательно увеличивать коэффициент α, стартуя с $\alpha = 1$. Применим процесс уточнения оценки запаса управляемости.) В результате получим, что при $\alpha = 1.037$ интервальная матрица $[Q^1(4,3)]$ – полного ранга, а при $\alpha = 1.038$ не полного ранга. Таким образом, в качестве оценки запаса управляемости ансамбля в окне (3,4) может быть взято число $\alpha = 1.037$ с точностью до трех знаков после запятой.

Заметим, что в окне (4,5) интервальная матрица $[Q^1(5,4)]$ в этом примере уже становится матрицей неполного ранга.

Заключение

Интервальную систему можно также воспринимать как частный случай континуума систем с переключением, когда применяются любые законы переключения динамических систем с одного уравнения на другое уравнение. Значит, утверждения теорем 1,2 могут быть применены для этого частного случая.

Библиографические ссылки

1. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. М.: Мир, 1987. 360 с.
2. Гайшун И.В. Системы с дискретным временем. Мн.: Институт математики НАН Беларуси, 2001. 400 с.
3. Шарый С.П. Об интервальных матрицах полного ранга // Сиб. журн. вычисл. математики / РАН Сиб. отделение. Новосибирск. 2014. Т. 17, № 3. С. 289–304.
4. Гантмахер Ф. Р. Теория матриц. М.: Наука, 1967. 576 с.

УДК 517.977.5

СТРАТЕГИЯ С ЗАМЫКАНИЕМ В ЗАДАЧЕ ОПТИМАЛЬНОГО ГАРАНТИРОВАННОГО УПРАВЛЕНИЯ И ЕЕ ПРИМЕНЕНИЕ В МРС

Д.А. Костюкевич, Н.М. Дмитрук

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, dmitrukn@bsu.by, kostukDA@bsu.by

Рассматривается задача оптимального управления линейной дискретной системой с неизвестными ограниченными возмущениями, которую требуется за конечное время перевести с гарантией на терминальное множество, обеспечивая при этом минимум гарантированного значения заданного критерия качества. Определяется оптимальная стратегия управления, учитывающая информацию об одном будущем состоянии объекта. Предлагается эффективный метод ее построения и алгоритм управления по прогнозирующей модели на основе оптимальных стратегий.

Ключевые слова: оптимальное гарантированное управление; стратегия управления; управление по прогнозирующей модели.

CLOSED-LOOP CONTROL STRATEGY FOR OPTIMAL GUARANTEED CONTROL PROBLEM AND ITS APPLICATIONS IN MPC

D.A. Kastsiukevich, N.M. Dmitruk

Belarusian State University, 4 Nezavisimosti Avenue., Minsk 220030, Belarus, dmitruknm@bsu.by, kostukDA@bsu.by

We consider an optimal control problem for a linear discrete-time system subject to unknown bounded disturbances, where the goal is to robustly steer its state to a certain terminal set in a finite time while minimizing guaranteed value of the cost function. We define optimal control strategy that takes into account information about one future state of the control object. We also suggest an effective design method for it and an algorithm for model predictive control based on it.

Keywords: optimal guaranteed control; control strategy; model predictive control.

Введение

Задачи оптимального управления системами, подверженными действию неизвестных возмущений или содержащими другие недоступные для непосредственных измерений параметры, для которых требуется получить гарантированный результат, рассматриваются в литературе с конца 60-х [1,2]. В настоящее время обширной областью применения задач оп-

тимального управления является теория управления по прогнозирующей модели (Model Predictive Control – MPC), а для задач управления в условиях неопределенности – ее робастная версия [3,4].

Теория MPC [3] опирается на решение в ходе конкретного процесса управления в каждый момент времени прогнозирующих задач оптимального управления на конечном промежутке времени с начальным состоянием, совпадающим с текущим состоянием процесса. Оптимальная программа прогнозирующей задачи подается на вход системы управления до тех пор, пока не будет получено и обработано следующее состояние. Повторяемая в темпе поступления информации о состоянии системы процедура дает реализацию обратной связи, обеспечивающую различные полезные свойства замкнутой системы. В робастном MPC, начиная с работы [4], для формирования обратной связи предлагается использовать не оптимальные гарантирующие программы, а оптимальные стратегии, учитывающие зависимость управляющих воздействий от информации о текущих и будущих состояниях процесса управления. Основное внимание уделяется определению стратегий, допускающих их эффективное построение [5].

Настоящее сообщение примыкает к работам [6–8], в которых стратегия управления определяется в предположении о коррекции управляющих воздействий в один будущий момент времени. В отличие от [6–8] качество управления оценивается функционалом из [5], что позволяет применять предложенные стратегии в MPC, определяет актуальность исследования и широкий спектр приложения результатов.

1. Постановка задачи и методология исследования

Рассмотрим дискретную линейную стационарную систему

$$x(t+1) = Ax(t) + Bu(t) + Mw(t),\ x(0) = x_0,\ t = 0,1,\ldots,T-1. \qquad (1)$$

Здесь $x(t) \in \mathbb{R}^n$ — состояние, $u(t) \in U \subset \mathbb{R}^r$ – управление, $w(t) \in W \subset \mathbb{R}^p$ —возмущение; $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times r}$, $M \in \mathbb{R}^{n\times p}$; $U = \{u \in \mathbb{R}^r : \| u \|_\infty \le u_{\max}\}$, $W = \{w \in \mathbb{R}^p : \| w \|_\infty \le w_{\max}\}$, где норма $\| z \|_\infty = \max_i | z_i |$.

Целью управления системой (1) является: 1) ее перевод с гарантией на терминальное множество $X_T = \{x \in \mathbb{R}^n : Hx \le g\}$; 2) минимизация гарантированного значения критерия качества как в [5]:

$$J(u) = \max_w \sum_{t=0}^{T-1} \left(\| Qx(t) \|_\infty + \| Ru(t) \|_\infty \right) + \| Px(T) \|_\infty . \qquad (2)$$

Здесь $H \in \mathbb{R}^{m\times n}$, $g \in \mathbb{R}^{m}$, $Q, P \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{r\times r}$. При построении MPC-регулятора матрицы Q, P, R являются параметрами настройки и подбираются в зависимости от желаемого поведения замкнутой системы [3].

Известно [4], что оптимальная гарантирующая программа $u^0(t) \in U$, $t = 0,\ldots, T-1$, приводит к консервативному поведению системы, поскольку не учитывает возможность поступления информации о ее поведении в будущем. Такую возможность учтем, определив далее стратегию управления.

Пусть $T_1 \in \{1,2,\ldots,T-1\}$ — некоторый момент времени. Он разбивает промежуток управления на $\Delta_0 = \{0,1,\ldots,T_1-1\}$ и $\Delta_1 = \{T_1, T_1+1,\ldots,T-1\}$. Следуя [6,8], назовем T_1 моментом замыкания системы (1). Для промежутков Δ_k, $k = 0,1$, определим:

$u_k(\cdot) = (u_k(t) \in U, t \in \Delta_k)$, $w_k(\cdot) = (w_k(t) \in U, t \in \Delta_k)$ − управление и возмущение, $U_k = \{u_k(\cdot) : u_k(t) \in U, t \in \Delta_k\}$, $W_k = \{w_k(\cdot) : w_k(t) \in W, t \in \Delta_k\}$ − множества доступных управлений и возможных возмущений на k-ом промежутке;

$X(T_1 \mid x_0, u_0(\cdot)) = \{x \in \mathbb{R}^{n} : x = x(T_1 \mid x_0, u_0(\cdot), w_0(\cdot)), w_0(\cdot) \in W_0\}$,

$X(T \mid x_1, u_1(\cdot)) = \{x \in \mathbb{R}^{n} : x = x(T \mid x_1, u_1(\cdot), w_1(\cdot)), w_1(\cdot) \in W_1\}$ − множества всех возможных состояний в моменты времени T_1 и T, где $x(t \mid x_k, u_k(\cdot), w_k(\cdot))$ − состояние системы (1) с начальным условием $x(0) = x_0$ или $x(T_1) = x_1$ под действием управления $u_k(\cdot)$ и возмущения $w_k(\cdot)$.

Предположение 1: До начала процесса управления известно, что в момент T_1 будет измерено состояние $x_1 = x(T_1 \mid x_0, u_0(\cdot), w_0(\cdot))$, с учетом которого выбирается новое управляющее воздействие $u_1(\cdot) = u_1(\cdot \mid x_1)$ на Δ_1.

С учетом предположения 1 и следуя работам [7,8], будем искать решение рассматриваемой задачи в виде *стратегии управления* с моментом замыкания T_1: $\pi_1 = \pi_1(0, x_0) = \{u_0(\cdot \mid x_0); u_1(\cdot \mid x_1), x_1 \in X(T_1 \mid x_0, u_0(\cdot \mid x_0))\}$, где $u(\cdot \mid x_k) = (u_k(t \mid x_k), t \in \Delta_k)$ — управляющее воздействие на Δ_k, $k = 0,1$.

Определение 1: Стратегия π_1 называется *допустимой стратегией управления с моментом замыкания* T_1, если выполняется включение $X(T \mid x_1, u_1(\cdot \mid x_1)) \subseteq X_T \ \ \forall x_1 \in X(T_1 \mid x_0, u_0(\cdot \mid x_0))$.

Определим оптимальную стратегию с моментом замыкания T_1

$$\pi_1^0 = \pi_1^0(0, x_0) = \{u_0^0(\cdot \mid x_0); u_1^0(\cdot \mid x_1), x_1 \in X(T_1 \mid x_0, u_0^0(\cdot \mid x_0))\}, \qquad (3)$$

где управление $u_0^0(\cdot \mid x_0)$ назовем оптимальной начальной программой.

Задача на промежутке Δ_1 для состояния x_1 заключается в отыскании оптимальной программы $u_1^0(\cdot \mid x_1)$, которая является решением задачи

$$J_1(x_1) = \min_{u_1(\cdot)\in U_1} \max_{w_1(\cdot)\in W_1} \Big\{ \sum_{t\in\Delta_1} \big(\| Qx(t \mid x_1, u_1(\cdot), w_1(\cdot)) \|_\infty + \| Ru_1(t) \|_\infty \big) + \quad (4)$$

$$+ \| Px(T \mid x_1, u_1(\cdot), w_1(\cdot)) \|_\infty \Big\},$$

при условии $x(T \mid x_1, u_1(\cdot \mid x_1), w_1(\cdot)) \in X_T \ \ \forall w_1(\cdot) \in W_1$.

Если задача (4) не имеет решения, полагаем $J_1(x_1) = +\infty$.

Предположение 2: Множество $X_1 = \{x_1 : J_1(x_1) < +\infty\}$ непусто, и $\exists u_0(\cdot \mid x_0)$, обеспечивающее включение $X(T_1 \mid x_0, u_0(\cdot \mid x_0)) \subseteq X_1$.

В предположении 2 стратегия $\pi_1 = \{u_0(\cdot \mid x_0); u_1^0(\cdot \mid x_1), x_1 \in X(T_1 \mid x_0, u_0(\cdot \mid x_0))\}$ допустима. Тогда оптимальная начальная программа $u_0^0(\cdot \mid x_0)$ на промежутке Δ_0 существует и является решением следующей минимаксной задачи

$$V(\pi_1^0) = \min_{u_0(\cdot)\in U_0} \max_{w_0(\cdot)\in W_0} \Big\{ \sum_{t\in\Delta_0} \big(\| Qx(t) \|_\infty + \| Ru_0(t) \|_\infty \big) + J_1(x(T_1)) \Big\}, \quad (5)$$

$$x(t+1) = Ax(t) + Bu_0(t) + Mw_0(t), \ \ x(0) = x_0, u_0(t) \in U_0, \ \ t \in \Delta_0,$$

$$x(T_1) \in X_1 \ \ \forall \ w_0(\cdot) \in W_0.$$

Определение 2: Стратегия управления (3) оптимальна, если $u_0^0(\cdot \mid x_0)$ – решение задачи (5), $u_1^0(\cdot \mid x_1)$ – решения задач (4) для состояний $x_1 \in X(T_1 \mid x_0, u_0^0(\cdot \mid x_0))$.

Для начала процесса управления необходимо знать лишь оптимальную начальную программу $u_0^0(\cdot \mid x_0)$. Оптимальные программы $u_1^0(\cdot \mid x_1)$ для промежутка Δ_1 заранее не строятся. В момент T_1, после измерения $x(T_1)$, вычисляется лишь одна из них – $u_1^0(\cdot \mid x(T_1))$. Поэтому цель дальнейшего изложения – эффективное вычисление оптимальной начальной программы.

2. Результаты и их обсуждение

Центральный результат настоящего исследования – простое описание множества $X_1(\alpha) = \{x_1 \in X_1 : J_1(x_1) \le \alpha\}$, позволяющее свести задачу (5) к задаче линейного программирования. Множество $X_1(\alpha)$ для дискретной системы (1) – многогранник. Пусть $p_i \in \mathbb{R}^n$, $i = 1, 2, \ldots, m_1$, $\| p_i \| = 1$, – нормали к граням $X_1(\alpha)$ и пусть

$$g_{p_i}(\alpha) = \max_{x_1 \in X_1(\alpha)} p_i' x_1. \qquad (6)$$

Легко установить, что (7) – задача линейного программирования, зависящая от параметра α. Тогда функция $g_{p_i}(\alpha)$ – вогнутая, кусочно-линейная. Для нее с использованием результатов параметрического линейного программирования может быть найдено разбиение отрезка $[\alpha_{\min}, \alpha_{\max}]$ допустимых значений параметра на области линейности функции $g_{p_i}(\alpha)$. Будем считать, что указанное разбиение найдено для всех p_i, $i = 1,\ldots,m_1$, в результате чего построено обобщающее разбиение $\alpha_{\min} = \alpha^1 < \ldots < \alpha^{K+1} = \alpha_{\max}$ и найдены значения $q_i^k = dg_{p_i}(\alpha^k + 0)/d\alpha$, $i = 1,\ldots,m_1$, $k = 1,\ldots,K$, причем для каждого i имеют место неравенства $q_i^1 > q_i^2 > \ldots > q_i^{K_i} \geq 0$.

Пусть $Q_1 \in \mathbb{R}^{m_1 \times K}$ – матрица, элементами которой являются найденные q_i^k, $P_1 \in \mathbb{R}^{m_1 \times n}$ – матрица, строками которой являются p_i, $i = 1,\ldots,m_1$, $g_1 = (g_{p_i}(\alpha_{\min}),\ i = 1,\ldots,m_1)$. Тогда $X_1(\alpha) = \{x \in \mathbb{R}^n : P_1 x \leq g_1 + Q_1\omega\}$, где вектор $\omega \in \mathbb{R}^K$ построен по правилам: $\omega_k = \alpha^{k+1} - \alpha^k$, $k = 1,\ldots,K(\alpha)-1$; $\omega_{K(\alpha)} = \alpha - \alpha^{K(\alpha)}$; $\omega_k = 0$, $k = K(\alpha)+1,\ldots,K$; $K(\alpha): \alpha \in [\alpha^{K(\alpha)}, \alpha^{K(\alpha)+1}[$.

Далее задача (5) для эффективного численного решения может быть сведена к задаче линейного программирования (см. например, [6,8]).

Отметим, что процедура построения матриц P_1, Q_1, значений α^k, может оказаться достаточно трудоемкой. В то же время этот этап является подготовительным и не зависит от x_0. Это означает, что в алгоритме МРС трудоемкость решения зависит только от трудоемкости решения задачи (5).

Приведем алгоритм МРС на основе оптимальных стратегий π_1^0. Прогнозирующая задача оптимального управления – это задача (5), в которой начальное состояние совпадает с текущим (в момент времени τ) состоянием объекта управления $x^*(\tau)$, т.е. начальное условие $x(0) = x_0$ заменяется на $x(0) = x^*(\tau)$. Оптимальная начальная программа прогнозирующей задачи, чтобы подчеркнуть момент τ, в который она получена, обозначается через $u_0^0(t \mid \tau)$, $t \in \Delta_0$.

Алгоритм МРС заключается в следующем:

1. в момент времени τ измерить $x^*(\tau)$ и найти $u_0^0(t \mid \tau)$, $t \in \Delta_0$;
2. применить к объекту управляющее воздействие $u_{MPC}(\tau) = u_0^0(0 \mid \tau)$;
3. перейти к следующему моменту ($\tau := \tau + 1$) и вернуться к шагу 1.

Численные эксперименты демонстрируют сравнимую трудоемкость построения оптимальных стратегий и оптимальных гарантирующих программ при улучшении качества процесса управления на стратегиях.

Библиографические ссылки

1. Куржанский А.Б. Управление и наблюдение в условиях неопределенности. М.: Наука, 1977.
2. Красовский Н.Н. Управление динамической системой. Задача о минимуме гарантированного результата. М.: Наука, 1985.
3. Rawlings J.B., Mayne D.Q. Model Predictive Control: Theory and Design. Madison: Nob Hill Publishing, 2009. 576 p.
4. Scokaert P.O.M., Mayne D.Q. Min-max feedback model predictive control for constrained linear systems // IEEE Trans. on Automatic control. 1998. Vol. 43, №. 8. P. 1136–1142.
5. Bemporad A., Borrelli F., Morari M. Min-max control of constrained uncertain discrete-time linear systems // IEEE Transactions on Automatic Control. 2003. Vol. 48, №. 9. P. 1600–1606.
6. Балашевич Н.В., Габасов Р., Кириллова Ф.М. Построение оптимальных обратных связей по математическим моделям с неопределенностью // Журнал вычислительной математики и математической физики. 2004. Т. 44, № 2. С. 265–286.
7. Kostyukova O., Kostina E. Robust optimal feedback for terminal linear-quadratic control
problems under disturbances // Math. programming. 2006. Vol. 107, №.1–2. P. 131–153.
8. Дмитрук Н.М. Оптимальная стратегия с одним моментом замыкания в линейной задаче оптимального гарантированного управления // Журнал вычислительной математики и математической физики. 2018. Т. 58, № 5. С. 664–681.

УДК 519.85

ИССЛЕДОВАНИЕ ПРОЦЕССОВ ОЦЕНКИ НЕОДНОРОДНОГО ПОТОКА В МУЛЬТИСЕТЯХ

Л.А. Пилипчук, М.П. Романчук

Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь, pilipchuk@bsu.by, fpm.romanchump1@bsu.by

Рассматривается задача минимизации размера множества обозреваемых узлов мультиграфа и локализации специальных программируемых устройств (сенсоров) в узлах с целью сбора необходимой информации о функции потока для оценки, управления и контроля трафика в той части сети, которая непосредственно не наблюдается. Исследование процессов моделирования мультипотока основано на конструктивной теории декомпозиции разреженных систем с использованием свойств разреженности матриц неполного и полного рангов.

Ключевые слова: Мультиграф, разреженная система; ранг; мультипоток; обозреваемый узел; сенсор; декомпозиция.

RESEARCH ON INHOMOGENEOUS FLOW ESTIMATION PROCESSES IN MULTINETWORKS

L.A. Pilipchuk, M.P. Romanchuk

Belarusian State University, 4 Niezalieznasci Avenue, Minsk 220030, Belarus, pilipchuk@bsu.by, fpm.romanchump1@bsu.by

The problem of minimizing the size of the set of monitored multigraph nodes and localization of special programming devices (sensors) in the nodes to collect the necessary information about the flow function to estimate, control and monitor the traffic in the part of the network that is not directly observed is considered. The study of multinetwork modeling processes is based on the constructive theory of decomposition of sparse systems using the sparsity properties of incomplete and full rank matrices.

***Keywords*:** Multigraph; sparse system; rank; multiflow; monitored node; sensor; decomposition.

Введение

В случае моделирования процесса оценки транспортного потока в масштабах крупных городов количество транспортных средств может достигать десятков тысяч, и изменения величин дугового и внешнего мультипотока обычно лежат в широком диапазоне. Стратегии полного перебора узлов сети с целью минимизации мощности множества обозреваемых узлов потребуют огромных вычислительных затрат. Важным яв-

ляется построение алгоритмических, структурных и технологических решений независимых подсистем с матрицами неполного/полного рангов с различными типами разреженности в синтезе с современными достижениями в области инновационных технологий разреженного матричного и сетевого анализа, алгоритмической теории графов, теоретической информатики.

1. Моделирование процессов оценки мультипотока

Рассмотрим конечный связный ориентированный мультиграф (мультисеть) $G = (I, U)$, где I – множество узлов, U – множество мультидуг, определенных на $I \times I$, $|I| < \infty$, $|U| < \infty$. Пусть $K = \{1, 2, \dots, r\}$ – множество, состоящее из r типов потока в мультисети G. Обозначим $G^k = (I^k, U^k)$ связную сеть, соответствующую типу потока $k \in K$, $I^k \subseteq I$, U^k – множество дуг для потока типа k, $k \in K$. Мы предполагаем, что мультиграф G двунаправленный, следовательно, если $\exists\, (i,j)^k \in U^k$, то $\exists\, (j,i)^k \in U^k, k \in K$.

Мы представляем трафик в мультисети $G = (I, U)$ функцией сетевого потока $x: U \to \mathbb{R}$, которая удовлетворяет следующей системе линейных алгебраических уравнений:

$$\sum_{\mathrm{j} \in I_i^+(U^k)} x_{ij}^k - \sum_{\mathrm{j} \in I_i^-(U^k)} x_{ji}^k = \begin{cases} x_i^k, & i \in I_k^*, \\ 0, i \in I^k \setminus I_k^*, k \in K, \end{cases} \quad (1)$$

где I_k^* – множество узлов с неизвестным внешним потоком x_i^k в узле $i \in I_k^*$, $I_k^* \subseteq I^k$.

Согласно [1, 2], если $I_k^* \neq \emptyset$, то ранг матрицы системы (1) для графа $G^k = (I^k, U^k)$ равен $|I^k|\ \forall\, k \in K$. Если $I_k^* = \emptyset$, то ранг матрицы системы (1) равен $\left|I^k\right| - 1$ (матрица инцидентности графа G^k), $k \in K$ [3].

Для того, чтобы получить информацию о неизвестных дуговых потоках x_{ij}^k для дуг $(i,j)^k \in U^k$ и внешних потоках x_i^k узлов $i \in I_k^*$, $k \in K$ сенсоры установлены в узлах мультиграфа G. Узлы мультиграфа с сенсорами назовем обозреваемыми и обозначим множество обозреваемых узлов мультиграфа $M = \bigcup_{k \in K} M_k$, где M_k – множество обозреваемых узлов для потока K. Множество M_k может быть пустым. Обозначим $K(i)$ – множество типов дуговых потоков, проходящих через узел $i \in I$. Мы предполагаем, что если узел i является обозреваемым, то известны значения дуговых потоков на всех исходящих и входящих дугах для узла $i \in M_k$, а также внешние потоки в узлах $i \in M_k \cap I_k^*\ \forall\, i \in M, \forall\, k \in K(i)$

$$\begin{gathered} x_{ij}^k = f_{ij}^k,\ \ j \in I_i^+(U^k),\ \ x_{ji}^k = f_{ji}^k, j \in I_i^-(U^k), \\ x_i^k = f_i^k, i \in M_k \cap I_k^* \neq \emptyset\ \ \forall\, i \in M, \forall\, k \in K(i), \end{gathered} \quad (2)$$

где $f_{ij}^k, f_{ji}^k, f_i^k$ – константы. Учтем дополнительную информацию о коэффициентах разбиения потока. Используя известные коэффициенты разбиения $p_{ij}^k,\ 0 < p_{ij}^k \leq 1, (i,j) \in U^k$, можно выразить общий исходящий из узла i поток $F(i) = \sum_{j \in I_i^+(U^k)} x_{ij}^k$ как функцию от объема потока по каждой исходящей дуге. На основании известных коэффициентов разбиения p_{ij}^k, $(i,j) \in U^k$ для дуг сети $G^k = (I^k, U^k), k \in K$, сформируем дополнительные уравнения взаимосвязи дуговых потоков следующим образом. Если для узла $i \in I$ выполняется соотношение $\left|I_i^+(U^k)\right| \geq 2$, то для любой дуги, исходящей из узла i, например $(i, v_i)^k$, дуговые потоки для всех дуг за исключением дуги $(i, v_i)^k$, исходящих из узла i, выразим через дуговой поток x_{i,v_i}^k следующим образом:

$$x_{ij}^k = \frac{p_{ij}^k}{p_{i,v_i}^k} x_{i,v_i}^k, \qquad j \in I_i^+(U^k) \setminus \{v_i\}, |I_i^+(U)| \geq 2, k \in K(i), \tag{3}$$

где $(i, v_i)^k$ – каноническая дуга, исходящая из узла i. Обозначим $\beta_{ij}^k = \frac{p_{ij}^k}{p_{i,v_i}^k}$, если $\left|I_i^+(U^k)\right| \geq 2$, и $\beta_{ij}^k = p_{i,v_i}^k$, если $\left|I_i^+(U^k)\right| = 1$. Определим численные значения дуговых потоков x_{ij}^k, которые можно выразить с помощью коэффициентов разбиения p_{ij}^k, p_{i,v_i}^k, через наблюдаемые потоки $x_{i,v_i}^k = f_{i,v_i}^k$, полученные от специальных программируемых устройств (сенсоров), установленных в узлах мультиграфа G (множество M). Численные значения дуговых потоков, выраженные с помощью (3) через коэффициенты разбиения потока, подставим в уравнения системы (1). Удалим из мультиграфа $G = (I, U)$ дуги с известными значениями дуговых потоков и узлы с известными значениями внешних потоков $x_i^k = f_i^k, k \in K(i),\ i \in M$. Исключим из системы (1) те уравнения, которые не содержат неизвестных дуговых и внешних потоков. Пусть q – число уравнений вида (3) с неизвестными значениями дуговых потоков. С учетом выполненных преобразований для мультиграфа G получим новый мультиграф $\bar{G} = (\bar{I}, \bar{U})$. В результате система (1), (3) относительно мультиграфа G преобразуется к виду

$$\begin{gathered} \sum_{j \in I_i^+(\bar{U}^k)} x_{ij}^k - \sum_{j \in I_i^-(\bar{U}^k)} x_{ji}^k = \begin{cases} a_i^k + x_i^k, & i \in \overline{I_k^*}, \\ a_i^k, & i \in \overline{I^k} \setminus \overline{I_k^*}, k \in K, \end{cases} \\ \sum_{(i,j) \in \bar{U}} \lambda_{ij}^{kp} x_{ij}^k = 0, p \in P = \{1, \dots, q\}, \text{если } P \neq \emptyset, \end{gathered} \tag{4}$$

$\overline{I_k^*}$ – множество узлов с неизвестным внешним потоком x_i^k в мультисети $\bar{G}$.

Сформулируем задачу размещения сенсоров для мультиграфа.

Найти минимальное число обозреваемых узлов M, для которого система (4) имеет единственное решение и определить узлы мультисети G для размещения сенсоров.

Результатам исследования и оценки однородного потока в двунаправленной сети посвящены работы [4–10].

2. Примеры локализации сенсоров в узлах мультиграфа

Пример 1. Для мультиграфа G, представленного на рисунке 1, установим сенсор в узел $M = \{5\}$, где $K(5) = \{2, 3\}$ – типы потока для узла $i = 5$. На рисунке 2 представлен мультиграф $G' = (I', U')$ после преобразования (2). Типы линий на рисунках 1 – 4 обозначены следующим образом: дуговой поток первого типа ($k = 1$) – сплошная линия, дуговой поток второго типа ($k = 2$) – прерывистая линия, дуговой поток третьего типа ($k = 3$) – пунктирная линия.

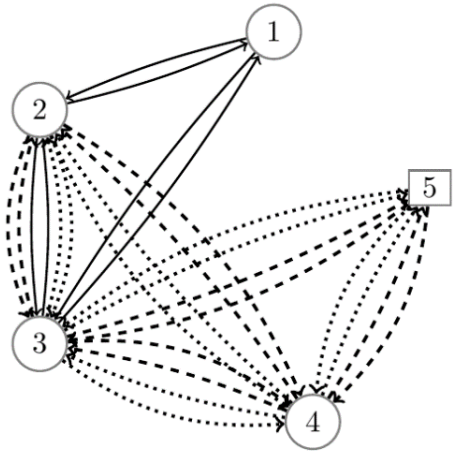


Рисунок 1 - Мультиграф $G = (I, U)$. Узел $M = \{5\}$ является обозреваемым

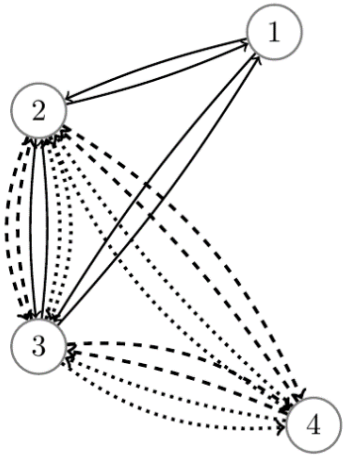


Рисунок 2 - Мультиграф $G' = (I', U')$ после преобразования (2)

Для узла 5 (пример 1) имеем $K(5) = \{2,3\}$. Поток типа $k = 1$ не входит в множество $K(5)$. Система (4) для $k = 1$ является недоопределенной.

Пример 2. Для мультиграфа G (рисунок 3) обозреваемым узлом M является узел 3, $M = \{3\}$. Мультиграф $G' = (I', U')$ после преобразования (2) представлен на рисунке 4.

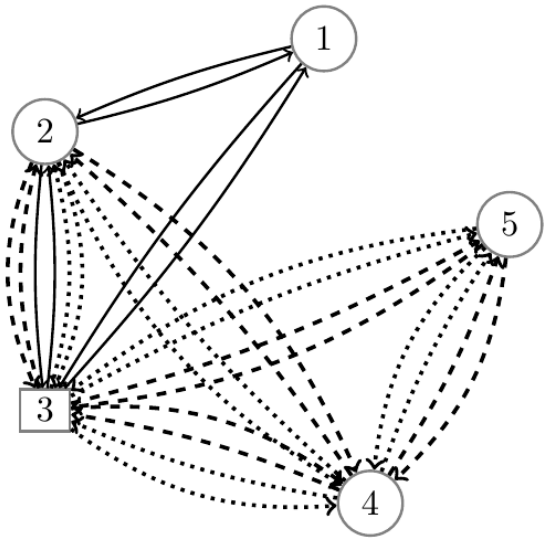


Рисунок 3 - Мультиграф $G = (I, U)$. Узел $M = \{3\}$ является обозреваемым

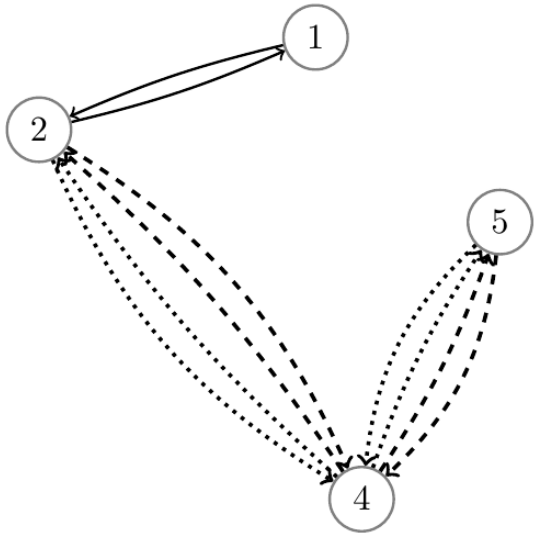


Рисунок 4 - Мультиграф $G' = (I', U')$ после преобразования (2)

Для узла $M = \{3\}$ система (4) для графа $\bar{G}$ с неизвестными дуговыми и внешними потоками имеет единственное решение.

Заключение

Предложен подход к созданию методов, алгоритмов и технологий решения задачи локализации специальных программируемых устройств (сенсоров) в узлах двунаправленного мультиграфа для оценки неоднородного потока на его ненаблюдаемых частях.

Библиографические ссылки

1. Пилипчук Л.А. О методах декомпозиции разреженных недоопределенных систем с матрицами полного и неполного ранга // Известия Гомельского государственного университета имени Ф. Скорины. 2016. № 6. С. 87–91.
2. Pilipchuk L.A., German O.V., Pilipchuk A.S. The general solutions of sparse systems with rectangular matrices in the problem of sensors optimal location in the nodes of a generalized graph // Вестник БГУ. Серия 1, Физика. Математика. Информатика. 2015. №2. С. 91–96.
3. Pilipchuk L.A. Sparse Linear Systems and Their Applications. Minsk: BSU, 2013. 236 p.
4. Bianco L, Confessore G, Gentili M. Combinatorial aspects of the sensor location problem. Annals of Operation Research. 2006. № 144(1). С. 201–234.
5. Bianco L, Confessore G, Reverberi P. A network based model for traffic sensor location with implication in O/D matrix estimates. Transportation Science. 2001. № 35(1). С. 50−60.

6. Bianco L, Cerrone C, Cerulli R, Gentili M. Locating sensors to observe network arc flows: exact and heuristic approaches. Computers and Operation Research. 2014. № 46. C. 12–22.
7. Gabasov R, Kirillova F.M, Kostyukova O.I. Конструктивные методы оптимизации. Часть 3. Сетевые задачи. Minsk: Belarusian State University; 1986.
8. Ahuja R.K., Magnanti T.L., Orlin J.B. Network flows: Theory, Algorithms, and Applications. New Jersey. 1993. 864 p.
9. Jensen P., Barnes D. Потоковое программирование. Москва: МГУ, 1984. 392 с.
10. Габасов Р., Кириллова Ф.М. Методы линейного программирования: в 3 частях. Ч. 3: Специальные задачи. Минск: БГУ, 1980. 368 с.