ISSN 2078-9181

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Отделение нанотехнологий и информационных технологий

САНКТ-ПЕТЕРБУРГСКИЙ
ИНСТИТУТ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ РАН

ТРУДЫ СПИИРАН

proceedings.spiiras.nw.ru

ВЫПУСК 6(43)

Санкт-Петербург
2015

18+

Труды СПИИРАН

Выпуск № 6(43), 2015

Научный, научно-образовательный, междисциплинарный журнал с базовой специализацией в области информатики, автоматизации и прикладной математики

Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН)

Адрес редакции

199178, Санкт-Петербург, 14-я линия, д. 39,
e-mail: publ@iias.spb.su, сайт: http://www.proceedings.spiiras.nw.ru/

Подписано к печати 15.12.2015. Формат 60×90 1/16. Усл. печ. л. 15,0. Заказ № 466. Тираж 150 экз., цена свободная
Отпечатано в Редакционно-издательском центре ГУАП, 190000, Санкт-Петербург, Б. Морская, д. 67

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций, свидетельство ПИ № ФС77-41695 от 19 августа 2010 г.
Подписной индекс 29393 по каталогу «Почта России»

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

SPIIRAS Proceedings

Issue № 6(43), 2015

Scientific, educational, and interdisciplinary journal primarily specialized
in computer science, automation, and applied mathematics

Trudy SPIIRAN ♦ Founded in 2002 ♦ Труды СПИИРАН

Founder and Publisher

Federal State Budget Institution of Science
St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences
(SPIIRAS)

Editorial Board's address

14-th line VO, 39, SPIIRAS, St. Petersburg, 199178, Russia,
e-mail: publ@iias.spb.su, web: http://www.proceedings.spiiras.nw.ru/

Signed to print 15.12.2015
Printed in Publishing center GUAP, 67, B. Morskaya, St. Petersburg, 190000, Russia

The journal is registered in Russian Federal Agency for Communications and Mass-Media Supervision,
certificate ПИ № ФС77-41695 dated August 19, 2010 г.
Subscription Index 29393, Russian Post Catalog

СОДЕРЖАНИЕ

CONTENTS

УДК 510.66

И.А. Рябинин, А.В. Струков

ПРЕДИСЛОВИЕ И ВСТУПИТЕЛЬНАЯ СТАТЬЯ К ПЕРЕИЗДАНИЮ РАБОТЫ П.С. ПОРЕЦКОГО «РЕШЕНИЕ ОБЩЕЙ ЗАДАЧИ ТЕОРИИ ВЕРОЯТНОСТЕЙ ПРИ ПОМОЩИ МАТЕМАТИЧЕСКОЙ ЛОГИКИ»

Рябинин И.А., Струков А.В. **Предисловие и вступительная статья к переизданию работы П.С.Порецкого «Решение общей задачи теории вероятностей при помощи математической логики».**

Аннотация. Предисловие и вступительная статья представляют переиздание работы Платона Сергеевича Порецкого, которая была записана как лекция 25 октября 1886 г. В предисловии дана краткая историческая справка о работах П.С.Порецкого в области математической логики и ее применимости к другим областям науки, в том числе и к теории вероятностей. Вступительная статья имеет основной целью показать, как в конце XIX века было сформировано начало логико-вероятностного анализа (ЛВА), суть которого состоит в корректном переходе от логического равенства между событиями к алгебраическому равенству между их вероятностями. Показано, что дальнейшее развитие ЛВА было вызвано практической потребностью в 60-х годах прошлого столетия в оценке надежности цифровых схем, а также надежности и безопасности структурно сложных систем. Обсуждается сложный математический и философский вопрос о сущности принципиально разных понятий – вероятностной логики (ВЛ) и логики вероятностей (ЛВ).

Ключевые слова: математическая логика, теория вероятностей, логико-вероятностный анализ, ортогональная дизъюнктивная нормальная форма (ОДНФ), функции алгебры логики (ФАЛ).

Ryabinin I.A., Strukov A.V. **A Preface and an Introductory Article to the Re-edition of the Work of Platon Sergeevich Poreckiy «Solving General Tasks in Probability Theory by Using Mathematical Logic».**

Abstract: A preface and introduction article presents an article by Platon Sergeevich Poreckij which is a record of his lecture delivered on October 25, 1886. The preface contains short historical reference about P. S. Poreckij's works in the field of mathematical logic and its application to other science, including the probability theory. The introduction article has the main goal to show how the beginning of logic-and-probabilistic method (LPM) was created at the end of the XIX century. LPM essence was in valid transition from logic equation between the events to algebraic equality between their probabilities. The article shows that LPM further development is connected to the necessity of evaluation of digital circuits reliability as well as structurally complex systems reliability and safety in 1960s. Scientific disputes and the possibility of combining mathematical logic and the probability theory do not stop in the XIX century. There are regular seminars and conferences held on this subject. We discuss the complex mathematical and philosophical question about the nature of fundamentally different concepts - the probabilistic logic (PL) and the logic of probability (LP).

Keywords: mathematical logic, probability theory, logical probabilities analysis, orthogonalization disjunctive normal form (ODNF), Boolean function (BF).

1. Предисловие. Переиздание классических работ выдающихся ученых в последние годы стало хорошей традицией. В научный оборот возвращаются труды, где впервые были введены некоторые важные

понятия, теории, методы и алгоритмы. В этой связи необходимо отметить, что в серии публикаций «Reprint from the Early Days of Information Sciences» международной лаборатории цифровой обработки сигналов университета г. Тампере (TICSP) под редакцией профессора Радомира С. Станковича (Radomir S. Stankoviĉ) и профессора Яакко Т. Астола (Jaakko T. Astola) в 2009 г. вышел репринт статьи П.С. Порецкого «Решение общей задачи теории вероятностей при помощи математической логики» и ее перевод на английский язык (http://ticsp.cs.tut.fi/reports/reprint-poreckij-r.pdf). В предисловии к репринту редакторы так объяснили цель публикации: «Исторические исследования научной дисциплины, как правило, есть признак её зрелости. При правильном проведении таких работ, этот вид исследований более чем перечисление фактов или предоставление кредита на определенные важные исследования. Это анализ путей мышления, которые привели к важным открытиям».

Хронологически первой работой великого русского ученого П.С. Порецкого (1846-1907), посвященной математической логике, является сообщение, читанное в 3-м заседании математической секции Общества Естествоиспытателей при Императорском Казанском университете астрономом-наблюдателем университета приват-доцентом П.С. Порецким 17 мая 1880г. и изданное в Собрании протоколов секции в 1881г [1].

Часто цитируемая, но не ставшая от того менее значимой, формулировка: «Математическая логика по предмету своему есть логика, а по методу математика» приведена П.С.Порецким в предисловии к сообщениям, прочитанным на заседаниях секции 27 февраля и 23 марта 1882 г. и изданным отдельным оттиском в 1884г. [2]. Именно тогда великий русский ученый в предисловии «Об отношении математической логики к математике и логике» отмечал, что все соглашаются, что математическая логика есть логика, а то, «что её метод вполне аналогичен *математическому* методу *алгебры* и ни в каком отношении ему не уступает,… это, конечно, требует доказательства».

Рассуждая о количественных формах алгебры и качественных формах логики, П.С. Порецкий замечает, что прямое перенесение, т.е. непосредственное применение принципов и приёмов алгебры к предмету логики *невозможно*, однако вполне возможно «…*приспособление* этих приёмов (с полным сохранением всей их *точности*) к изучению качественных форм…» [2].

Наиболее важное критическое замечание П.С. Порецкого относится именно к гипотезе Дж. Буля [4] о тесной связи между

алгеброй и логикой, «…в силу которой при известных условиях (которые Буль указывает, но повторять которые здесь было бы вполне излишне), формулы и приемы алгебры могут быть переносимы в логику, и обратно. Эта гипотеза столь невероятна (смешивает свойства количества и качества), что подрывает всякое доверие к способу и вообще ко всей логической системе Буля. Независимо от этого, в способе Буля очень странно действует на читателя чередование логических приёмов с математическими и невозможность дать себе отчет в том, какие процессы мысли отвечают различным фазисам применяемого метода. Благодаря этому обстоятельству, доступны пониманию только первоначальное равенство и окончательный результат; все же остальное загадочно и произвольно…» [2].

Далее для описания духа критики работы Дж. Буля следует привести пусть пространную, но весьма значимую цитату работы П.С. Порецкого:

«…В настоящее время способ Буля (да и вообще все его учение) может представлять только исторический интерес, и мы привели его лишь затем, чтобы засвидетельствовать дань уважения глубокому уму, который, не имея предшественников (в сколько-нибудь серьезном смысле этого слова), положил прочное основание новой отрасли знаний, установив целый ряд бесспорных положений (независящих от упомянутой гипотезы) и указав задачи, настолько трудные и сложные, что для решения их помимо гипотезы оказалось недостаточным всего его остроумия. К счастью, другой достойный математик, Шредер, уделив часть своего досуга вопросам логики, успел разобраться среди лабиринта идей и приемов Буля, отделил в его учении произвольное от доказанного, усовершенствовал обозначения и вид формул, которые оказалось возможным удержать, и таким образом сохранил для науки те истины, которые были открыты Булем при сооружении его, хотя и блестящего, но эфемерного здания математической логики….» [2].

Как отмечал сам П.С. Порецкий, на тот момент у него не было возможности высказаться о применении математической логики к другим областям науки, кроме теории умозаключений. И хотя именно в это время профессор Казанского университета А.В. Васильев «…доставил возможность иметь в своем распоряжении весьма редкое сочинение Буля (первого автора по математической логике)» [2], и П.С. Порецкий знал об идеи Дж. Буля и его учеников У.С. Джевонса и Э. Шредера применить математическую логику к теории вероятностей, теории статистических отношений, теории отношений причин к следствиям, но считал этот вопрос «совершенно открытым».

И только в 1886г. П.С. Порецкий в сообщении, читанном 25 октября на 60-м заседании секции физико-математических наук Общества Естествоиспытателей при Императорском Казанском университете высказывает свое мнение о попытке Дж.Буля решить общую задачу теории вероятностей методами математической логики. Следует отметить, что как таковая задача «*приспособления*» методов алгебры к «изучению качественных форм» у него не стояла, или, возможно Дж. Буль эту операция производил в уме.

Но именно на этой задаче «*приспособления*» и сосредоточил свое внимание П.С. Порецкий в работе 1887 [3], используя результаты, изложенные в работе 1884г. И, если в трудах Дж. Буля вопрос о *приспособлении* алгебраических методов к предметам логики решался не явно, можно сказать загадочно, то разгадку этого как раз и предложил П.С. Порецкий в работе 1887 года [3].

Для более точной оценки значимости работы П.С. Порецкого хотелось бы привести весьма точное высказывание известного специалиста по истории математической логики в России В.А. Бажанова, по работе [5] которого приведена библиография работ П.С. Порецкого: «Значительная заслуга П.С. Порецкого состоит в том, что математическая логика стала развиваться не в направлении решения уравнений и удаления неизвестных, а в направлении получения всевозможных следствий из данных посылок» [5].

Поэтому в программе первого в России курса по математической логике, который вел приват-доцент П.С. Порецкий, последний, заключительный раздел программы назывался «О вероятностях логических классов. Об определении вероятностей событий при помощи математической логики». Наверно, следует пожелать включения подобного раздела, например под названием «Логико-вероятностное исчисление» и в современные учебные программы по математической логике.

Копия работы П.С. Порецкого любезно предоставлена сотрудниками Библиотеки Российской академии наук в Санкт-Петербурге, положительно оценившими идею переиздания работы выдающегося русского ученого, оригинал которой хранится в библиотеке Казанского университета. При переиздании максимально сохранены особенности стиля, орфографии, пунктуации, вариантность сокращений. Внесены незначительные правки в слова, соответствующие современной стилистике.

И в заключении предисловия хотелось бы привести ответ сербского профессора Радомира С. Станковича на запрос о его мнении по поводу переиздания работы П.С.Порецкого: «Профессор J.T. Astola

и я очень рады узнать, что работа П.С. Порецкого будет опубликована в журнале. Мы считаем, что эта работа П. С. Порецкого является очень важным историческим вкладом в область вероятностного анализа при помощи логических методов».

2. Вступительная статья «Платон Сергеевич Порецкий (1846-1907) – первооткрыватель логико-вероятностного анализа». В §1 Сообщения [3] он ставит философский вопрос: возможно ли приложение учения о качественных символах (логических классах) к учению о символах количественных (вероятностных)? И отвечает: возможно.

Этот вопрос до сих пор ставит в тупик некоторых математиков [6,7]. Так профессор Голота Я.Я. считает «Алгебра логики высказываний исходит из полной определённости объектов изучения. Теория же вероятностей предполагает неопределённость в совершении событий. Таким образом, в одной теории объединяются отрицающие друг друга начала: полная определённость и неопределённость. Не говорит ли это об очевидном противоречии, лежащем в основе логико-вероятностной теории?».

Другой ученый доктор технических наук Соколюк В.Н. в работе [7, с.103], рассуждая об адекватности логики мировоззренческим принципам, пишет… «Если придерживаться теории вероятностей и алгебры логики в том виде, в каком они сложились на сегодня, то следует признать, что абсурдно говорить об «истинности событий» и о «вероятности высказываний», поскольку истинность – характеристика высказываний, но не событий, а вероятность – характеристика событий, но не высказываний. Каждое высказывание феноменально. Бессмысленно говорить об их массовости в теоретико-вероятностном смысле, хотя многие «ученые» даже пишут книги и создают теории, которые ведут в никуда [6]».

И только С.Н. Берштейн не испугался формальной разницы между качественными и количественными символами, и ровно через 30 лет разработал первую (по времени) аксиоматику логики высказываний для аксиоматизации теории вероятностей [8]. Думаю, что Сергей Натанович был знаком с работой П.С. Порецкого [3], который в 1870 году окончил физико-математический факультет Харьковского университета, в котором с 1907 по 1933 г. преподавал и С.Н. Берштейн.

Невостребованность практикой (первой половиной XX столетия) привела к забвению выдающихся математических результатов Порецкого П.С. и Берштейна С.Н. Этому способствовали и трудности в добыче их публикаций. Ссылаясь на работу [3] только

по ее названию, многие авторы не отождествляли Порецкого П.С. с первооткрывателем логико-вероятностного анализа.

Так на с.3 [3] дано первое определение ЛВА…«Отсюда открывается общий путь для определения вероятностей: найти логическую связь между событиями, которого вероятность ищется, и другими событиями, вероятности которых даны, а затем сделать *переход* от логического равенства между событиями к алгебраическому равенству между их вероятностями». Изюминка этого определения кроется в выделенном автором слове «*переход*». Отсюда начинается описание главного результата автора. Для возможности пользоваться правилом несовместности необходимо уметь каждый логический многочлен:

$$A \vee B \vee C \vee D \vee ... \qquad (1)$$

приводить к *дисъюнктному* (по современному – ортогональному) виду, т.е. к виду:

$$A \vee \overline{A}B \vee \overline{A}\overline{B}C \vee \overline{A}\overline{B}\overline{C}D \vee ..., \qquad (2)$$

где $\overline{A}$ есть отрицание *A,* $\overline{B}$ - отрицание *B* и т.д.

Здесь я привел современные правила обозначения логических сумм $\vee$ и отрицаний $\overline{A}$ (у П.С.Порецкого «+» и *Ao* соответственно).

Оба многочлена логически равнозначны, но отличаются тем, что к первому из них не применима теорема о вероятности суммы несовместных событий, тогда как к второму применимо. Вероятность:

$$P(A \vee B \vee C \vee D), \qquad (3)$$

будучи приведена к *дисъюнктному* виду, разбивается на сумму вероятностей:

$$P(A) + P(\overline{A}B) + P(\overline{A}\overline{B}C) + P(\overline{A}\overline{B}\overline{C}D). \qquad (4)$$

Из теории вероятностей известно, если два и более события суть независимы, то вероятность их совпадения равна произведению их отдельных вероятностей. Это значит, что если *a,b,c* … - суть простые события, не связанные между собою никакими логическим отношениями, то $P(abc...) = P(a)P(b)P(c)...$. Тогда (4) может записано в следующем виде:

$$P(A) + P(\overline{A})P(B) + P(\overline{A})P(\overline{B})P(C) + P(\overline{A})P(\overline{B})P(\overline{C})P(D). \qquad (5)$$

На стр. 7 [3] в качестве примера представлен современный алгоритм ортогонализации для дизъюнкции $ab \vee cd$:

1) Проводится внешний цикл ортогонализации: $ab \vee cd = ab \vee \overline{ab}cd$;

2) Затем отрицание $\overline{ab}$ по закону де Моргана преобразуется в дизъюнкцию двух отрицание: $\overline{ab} = \bar{a} \vee \bar{b}$;

3) Проводится внутренний цикл ортогонализации: $\bar{a} \vee \bar{b} = \bar{a} \vee a\bar{b}$;

4) Объединяются все три операции:

$$ab \vee cd = ab \vee \overline{ab}cd = ab \vee (\bar{a} \vee \bar{b})cd = ab \vee (\bar{a} \vee a\bar{b})cd = \\ = ab \vee \bar{a}cd \vee a\bar{b}cd. \qquad (6)$$

Выражение (6) есть ортогональная дизъюнктивная нормальная форма (ОДНФ), которая позволяет вычислить вероятность $P(ab \vee cd) = P(ab) + P(\bar{a}cd) + P(a\bar{b}cd)$.

Таким образом, именно Порецкий П.С. в 1886 году открыл строгий математический метод вычисления вероятности сложного события через вероятности простых событий, т.е. метод, который в 1963 году получил название логико-вероятностного метода (ЛВМ) [9]. Это вторичное независимое открытие алгоритма ортогонализации произошло в 1963 году в Институте математики (Новосибирск) в отделении Вычислительной техники специалистом по счетно-решающим приборам и устройствам Мерекиным Юрием Владимировичем. В это время задача о вероятности вычисления обращения в единицу булевой функции уже считалась тривиальным решением. Для решения прикладных задач применение совершенной нормальной дизъюнктивной формы (СДНФ) считалось нерациональным из-за большого числа дизъюнктивных членов. Возникла необходимость построения «короткой» ортогональной формы, которая и была получена в 1963 году [9].

Одно из первых описаний применения ЛВМ для задач оценки надежности структурно-сложных технических систем на примере расчета надежности судовых электро-энергетических систем (СЭС) относится к 1967 г. [10]. Суть ЛВМ сформулирована в следующем виде: «…Метод расчета надежности судовых электроэнергетических систем, при котором структура СЭС описывается средствами математической логики, а количественная оценка ее надежности производится с помощью теории вероятностей, будем называть

логико-вероятностным методом» [10, с.249]. Дальнейшее развитие ЛВМ в области надежности и безопасности структурно-сложных систем связано с созданием научной школы профессора И.А.Рябинина «Логико-вероятностные методы исследования надежности, живучести и безопасности структурно сложных систем».

В зарубежной литературе описание метода ортогонализации в задачах оценки надежности структурно сложных систем (на примере расчета надежности сети ARPA) появилось в 1973 году в работе итальянцев Luidge Fratta и Ugo Montanari [11]. Авторы отмечали в частности, что использование концепции Булевой алгебры возможно и во многих других отраслях науки, например, в теории переключающих устройств или теории кодирования. В обзоре [12] приведен список из 150 публикаций зарубежных периодических изданий, которые в той или иной мере имели отношению к использованию алгоритмов ЛВМ. В начале 21 века в связи с новыми результатами в области работ, связанных с искусственным интеллектом, особенно активно и широко обсуждаются вопросы разработки и применения алгоритмов Sum of Disjoint Products (SDP) – «сумма несовместных произведений», а в нашей терминологии – ОДНФ.

Эволюцию идей математической логики нельзя представить в виде восходящей кривой. Периоды расцвета зачастую сменяются моментами регрессии и частичного упадка. В связи с критикой ЛВА Я.Я. Голотой [6] и Соколюком В.Н. [7] полезно вспомнить критику московского логика и математика Б.М. Кояловича [13, стр.417], который огонь своей критики направлял против практической неэффективности алгебры логики, которая, как ему казалось, была принципиально не способна давать плодотворные внелогические приложения.

Переводчик на русский язык книги Л. Кутюра «Алгебра логики» профессор И.В. Слешинский, отвечая Кояловичу, смог лишь указать на работу Порецкого «Решение общей задачи теории вероятностей при помощи математической логики», заметив, что Порецкий несколько рационализировал и развил Буля в этом вопросе.

Тогда Коялович сослался на то обстоятельство, что ни в одном из крупных современных ему трактатов по теории вероятностей (и среди них в монографии русского академика А.А.Маркова) нет никакого упоминания о существовании какой-либо связи этой научной дисциплины с алгеброй логики.

Однако в 1910 году физик Пауль Эренфест первым предложил использовать математическую логику в технике. Он писал: «Символическая формулировка даст возможность «вычислять»

следствия из таких сложных посылок, в которых при словесном изложении почти или совершенно невозможно разобраться». В качестве примера он приводил схемы проводов автоматической телефонной станции.

Выяснение аналогии между математической логикой и теорией вероятностей имеет как теоретический, так и практический интерес. Проблематика связи логики с вероятностью начала развиваться в древности Аристотелем, затем Г.В. Лейбницем, Дж. Булем, У.С. Джевонсом, Дж. Венном, Р. Карнапом и другими. Не углубляясь в века, рассмотрим эту связь на уровне 19-20 веков, когда возникла математическая логика Буля и завершилось формирование современной теории вероятностей С.Н. Берштейном и А.Н. Коломогоровым.

Анализ взаимоотношений между вероятностью и логикой в междисциплинарном плане в наше время регулярно рассматривается на специальных семинарах в Великобритании.

Так на семинаре Огастеса де Моргана в королевском колледже, (Лондон 4-6 ноября 2002г.) обсуждались вопросы: - как вероятность относится к логике? – может ли объединяться вероятность и логика? – если да, то как? [14].

В специальном выпуске 3-го семинара в 2007 году анонсируется статья Колина Хаусана (Colin Howson) «Можно ли логику объединить с вероятностью?» [15]. В 2015 году (20-24 апреля) прошел 7-й семинар «Объединение вероятности и логики» в университете Кентберри [16].

Вероятностная логика возникла как непосредственное продолжение индуктивной логики. Значения истинности в вероятностной логике называются вероятностями истинности высказываний, степенями правдоподобия или подтверждения.

Логика вероятностей, в которой высказываниям приписываются исключительно значения истины и лжи как в двухзначной логике.

В настоящее время вероятностная логика находит наибольшее применение в развитии приложений к искусственному интеллекту [17], а логика вероятностей в середине 20 века нашла применение к решению проблем надежности, живучести и безопасности структурно-сложных систем [18, 19].

Смысл слов «вероятностная логика (ВЛ)» и «логика вероятностей (ЛВ)» долгое время *воспринимался как синонимы*.

А сущность этих принципиально разных понятий состоит в следующем:

– предметом *вероятностной логики* Д.М. Кейнса (Keynes J.M.) [20], Дж. Фон Неймана [21] и Нильса Нильссона (Nilsson N.J.) [17] является оценка истинности гипотез (высказываний), которые заключены в промежуток между «истиной» и «ложью» ($1 \geq x \geq 0$);

– предметом *логики вероятностей* Джорджа Буля [4], П.С. Порецкого [3], Рябинина И.А. [10, 18, 19] является вычисление вероятности истинности случайных событий (высказываний), принимающих только два значения (1;0).

В первом случае имеют дело с многозначной логикой, во втором – с двухзначной логикой.

Теории логики, допускающие более чем две категории «истинных» и «ложных» высказываний, составляют то, что обычно называют «модальной» логикой, а допускаемые ими категории - «модусами» или «степенями правдоподобия». Модальная логика оперирует такими истинностными значениями, как «возможно», «необходимо» и т.д.

Необходимость применения в логике вероятностных методов диктовалась прогрессом развития самой математической логики и теоретической информатики.

Диссертация Сперанского С.О. «Логика вероятностей и вероятностная логика» [22] посвящена изучению математической стороны обоих этих подходов. По утверждению Сперанского С.О.:

– цель вероятностной логики – введение в рассмотрение и дальнейшее изучение разнообразных языков для рассуждений о вероятностях.

– логика вероятностей ставит во главу угла проблему индуктивного синтеза непротиворечивых теорий.

Первооткрывателей логико-вероятностного анализа (Дж. Буля и П.С.Порецкого) в связи с отождествлением ЛВ и ВЛ практически все ученые считали сторонниками именно вероятностной логики, а не логики вероятностей.

3. Примеры. Чтобы практически осознать нестандартность вычисления вероятностей на сложных структурах и понять сущность логики вероятностей в работе [23] показан пример вычислений вероятностей функций алгебры логики (ФАЛ) структурно сложной системы (рисунок 1). Рассмотрим две из четырех ФАЛ указанного примера.

Пример №1. Анализируемая система состоит из двух антенн x_1 и x_2, переключающего устройства x_5 и двух приемных устройств x_3 и x_4.

Система работоспособна, если сигнал получен на выходе хотя бы одного приемника.

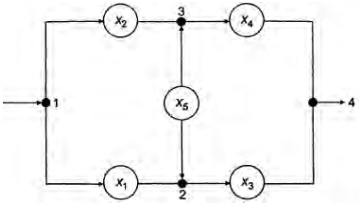


Рис. 1. Мостиковая схема

Логический критерий работоспособности системы может быть записан в виде дизъюнкции $Yc_1 = x_3 \vee x_4$. ФАЛ, соответствующая логическому критерию, может быть записана в матричной форме:

$$Y_{c1} = x_3 \vee x_4 = \begin{vmatrix} x_1 x_3 \\ x_2 x_4 \\ x_1 x_5 x_4 \\ x_2 x_5 x_3 \end{vmatrix} = \begin{vmatrix} K_1 \\ K_2 \\ K_3 \\ K_4 \end{vmatrix}, \qquad (7)$$

где конъюнкции $K_1 = x_1 x_3$, $K_2 = x_2 x_4$, $K_3 = x_1 x_5\, x_4$, $K_4 = x_2 x_5\, x_3$ образуют дизъюнктивную нормальную форму ФАЛ, которая в обычной записи имеет вид: $Yc_1 = K_1 \vee K_2 \vee K_3 \vee K_4$.

Чтобы определить вероятность:

$$P\{y_{c1}(x_1, x_2, ..., x_5) = 1\} \qquad (8)$$

при известных вероятностях истинности исходных высказываний:

$$P\{x_i = 1\} = R_i \;\; , P\{x_i = 0\} = Q_i \qquad (9)$$

категорически нельзя заменять качественные символы x_i количественными символами R_i, Q_i в ФАЛ (7) из-за повторности в них некоторых x_i и совместности конъюнкций $K_1 \div K_4$.

Еще Н. Руш в 1956г. [24] рекомендовал полный перебор всех возможных состояний системы путем записи функции алгебры логики (ФАЛ) в совершенной дизъюнктивной нормальной форме (СДНФ).

Учитывая громадное число возможных состояний 2^n в реальных задачах, когда n равно не 5, а несколько десятков и сотен, все разработчики логико-вероятностных методов (ЛВМ) искали

соответствующие алгоритмы преобразования самих ФАЛ, чтобы при сохранении абсолютной точности расчетов добиться существенного сокращения их трудоемкости.

Воспользуемся одним из них – алгоритмом ортогонализации [19].

Алгоритм ортогонализации в современных терминах условно можно представить как последовательное выполнение внешнего и внутреннего цикла. Такое представление реализовано в большинстве современных компьютерных программ. Внешний цикл ортогонализации позволяет представить ФАЛ как дизъюнкцию несовместных конъюнкций. В нашем примере процедура внешнего цикла преобразует (7) следующим образом:

$$Y_{c1} = x_3 \vee x_4 = \begin{vmatrix} K_1 \\ K_2 \\ K_3 \\ K_4 \end{vmatrix} = \begin{vmatrix} K_1 \\ \overline{K_1} K_2 \\ \overline{K_1}\,\overline{K_2} K_3 \\ \overline{K_1}\,\overline{K_2}\,\overline{K_3} K_4 \end{vmatrix}. \qquad (10)$$

В работе П.С.Порецкого [3] процедура внешнего цикла описана как приведение логического многочлена (1) к дизъюнктному виду, то есть к виду (2). Доказательство корректности такого преобразования дано П.С.Порецком в работе [1]. Необходимость такого преобразования объясняется возможностью применения теоремы о вероятности суммы несовместных событий. В этом случае равенство (8) с учетом (10) разбивается на сумму вероятностей:

$$\begin{aligned} &P\{y_{c1}(x_1, x_2, ..., x_5) = 1\} = P\{K_1 = 1\} + P\{\overline{K}_1 K_2 = 1\} + \\ &+ P\{\overline{K}_1 \overline{K}_2 K_3 = 1\} + P\{\overline{K}_1 \overline{K}_2 \overline{K}_3 K_4 = 1\}. \end{aligned} \qquad (11)$$

Нетрудно заметить, что в конъюнкциях $K_1 \div K_4$ все логические переменные встречаются дважды. Повторность логических переменных должна быть устранена при проведении внутреннего цикла ортогонализации.

Содержание внутреннего цикла ортогонализации состоит в нахождении отрицаний конъюнкций в ортогональной форме, то есть в виде:

$$\overline{x_i x_j x_k} = \overline{x}_i \vee \overline{x}_j \vee \overline{x}_k = \overline{x}_i \vee x_i \overline{x}_j \vee x_i x_j \overline{x}_k, \qquad (12)$$

и корректном перемножении конъюнкций. При этом, кроме правила (12) и закона тавтологии ($x_i x_i = x_i$) в современных алгоритмах применяются правила сокращенного умножения для инверсных конъюнкций, например, $\overline{AB} \cdot A = \overline{B} \cdot A$, $\overline{AB} \cdot \overline{CB} = \overline{B} \vee B\overline{AC}$ и т.д.

Для нашего примера формула (10) с учетом (7) преобразуется следующим образом:

$$Y_{c1} = \begin{vmatrix} K_1 \\ \overline{K_1} K_2 \\ \overline{K_1}\,\overline{K_2} K_3 \\ \overline{K_1}\,\overline{K_2}\,\overline{K_3} K_4 \end{vmatrix} = \begin{vmatrix} x_1 x_3 \\ \overline{x_1 x_3}\, x_2 x_4 \\ \overline{x_1 x_3}\;\overline{x_2 x_4}\; x_1 x_5 x_4 \\ \overline{x_1 x_3}\;\overline{x_2 x_4}\;\overline{x_1 x_5 x_4}\; x_2 x_5 x_3 \end{vmatrix} = \tag{13}$$

$$= \begin{vmatrix} x_1 x_3 \\ (\overline{x}_1 \vee x_1 \overline{x}_3)\; x_2 x_4 \\ \overline{x}_3 \overline{x}_2\; x_1 x_5 x_4 \\ \overline{x}_1 \overline{x}_4 (\overline{x}_1 \vee x_1 \overline{x}_4)\; x_2 x_5 x_3 \end{vmatrix} = \begin{vmatrix} x_1 x_3 \\ \overline{x}_1 x_2 x_4 \vee x_1 \overline{x}_3\; x_2 x_4 \\ \overline{x}_3 \overline{x}_2\; x_1 x_5 x_4 \\ \overline{x}_1 \overline{x}_4\; x_2 x_5 x_3 \end{vmatrix}.$$

После преобразования функции (7) в ортогональную дизъюнктивную форму (13) мы получили выражение, в котором возможно полное замещение x_i на R_i, Q_i, логических сумм и произведений на алгебраические.

Используя вероятности (9) ортогональную дизъюнктивную форму (13), вычислим вероятностную функцию (ВФ):

$$P\{y_{c1}(x_1, x_2, ..., x_5) = 1\} = \begin{matrix} R_1 R_3 + Q_1 R_2 R_4 + R_1 Q_3 R_2 R_4 + \\ + Q_3 Q_2 R_1 R_5 R_4 + Q_1 Q_4 R_2 R_5 R_4 \end{matrix}. \tag{14}$$

Пример №2. Пусть критерий опасного состояния будет (15) (см.с.93 [19]).

$$Y_{c2} = \begin{vmatrix} x_1 x_3 x_4 \\ x_1 x_3 x_5 \\ x_2 x_4 x_3 \\ x_2 x_5 x_4 \end{vmatrix}. \tag{15}$$

Результатом проведения внешнего цикла ортогонализации матрицы (15) будет матрица вида:

$$Y_{c2} = \begin{vmatrix} \overline{x_1 x_3 x_4} \\ \overline{x_1 x_3 x_4}\, \overline{x_1 x_3 x_5} \\ \overline{x_1 x_3 x_4}\, \overline{x_1 x_3 x_5}\, \overline{x_2 x_4 x_3 x_2 x_5 x_4} \end{vmatrix}. \qquad (16)$$

После применения в (16) теоремы де Моргана и замены отрицаний конъюнкций на сумму отрицаний получим:

$$Y_{c2} = \begin{vmatrix} x_1 x_3 x_4 \\ \begin{vmatrix} \overline{x}_1 \\ \overline{x}_3 \\ \overline{x}_4 \end{vmatrix} x_1 x_3 x_5 \\ \begin{vmatrix} \overline{x}_1 & \begin{vmatrix} \overline{x}_1 \\ \overline{x}_3 \\ \overline{x}_5 \end{vmatrix} \\ \overline{x}_3 & \\ \overline{x}_4 & \end{vmatrix} x_2 x_4 x_3 \\ \begin{vmatrix} \overline{x}_1 & \begin{vmatrix} \overline{x}_1 & \begin{vmatrix} \overline{x}_2 \\ \overline{x}_4 \\ \overline{x}_3 \end{vmatrix} \\ \overline{x}_3 & \\ \overline{x}_4 & \end{vmatrix} \\ \overline{x}_3 & \\ \overline{x}_4 & \end{vmatrix} x_2 x_5 x_4 \end{vmatrix}. \qquad (17)$$

Результатом проведения внутреннего цикла ортогонализации матрицы (17) будет матрица вида:

$$Y_{c2} = \begin{vmatrix} x_1 x_3 x_4 \\ \begin{vmatrix} \overline{x}_1 \\ x_1 \overline{x}_3 \\ x_1 x_3 \overline{x}_4 \end{vmatrix} x_1 x_3 x_5 \\ \begin{vmatrix} \overline{x}_1 & \begin{vmatrix} \overline{x}_1 \\ x_1 \overline{x}_3 \\ x_1 x_3 \overline{x}_5 \end{vmatrix} \\ x_1 \overline{x}_3 & \\ x_1 x_3 \overline{x}_4 & \end{vmatrix} x_2 x_4 x_3 \\ \begin{vmatrix} \overline{x}_1 & \begin{vmatrix} \overline{x}_1 & \begin{vmatrix} \overline{x}_2 \\ x_2 \overline{x}_4 \\ x_2 x_4 \overline{x}_3 \end{vmatrix} \\ x_1 \overline{x}_3 & \\ x_1 x_3 \overline{x}_5 & \end{vmatrix} \\ x_1 \overline{x}_3 & \\ x_1 x_3 \overline{x}_4 & \end{vmatrix} x_2 x_5 x_4 \end{vmatrix} = \begin{vmatrix} x_1 x_3 x_4 \\ x_1 x_3 \overline{x}_4 x_5 \\ \overline{x}_1 x_2 x_4 x_3 \\ \overline{x}_1 x_2 \overline{x}_3 x_5 x_4 \\ x_1 x_2 \overline{x}_3 x_5 x_4 \end{vmatrix} = \begin{vmatrix} x_1 x_3 x_4 \\ x_1 x_3 \overline{x}_4 x_5 \\ \overline{x}_1 x_2 x_4 x_3 \\ x_2 \overline{x}_3 x_5 x_4 \end{vmatrix}. \qquad (18)$$

Используя вероятности (9) ортогональную дизъюнктивную форму (18), вычислим вероятностную функцию (ВФ):

$$P\{y_4(x_1, x_2, ..., x_5) = 1\} = R_1R_3R_4 + R_1R_3Q_4R_5 + Q_1R_2R_4R_3 + R_2Q_3R_5R_4 . \quad (19)$$

4. Заключение. В § 42 [25] «Вероятностные методы в логике» говорится, что вероятностные методы исследования введены в формальную логику сравнительно недавно. Необходимость применения в логике вероятностных методов сохраняется и даже делается более настоятельной. Сообщается, что еще Джордж Буль предложил логическую интерпретацию частотной (статистической) вероятности как вероятность суждений о событиях. На стр.198 приведены такие слова: «Буль исследовал связь «алгебры логики» с обычной алгеброй и интерпретировал не только как логику классов (терминов), но и как логику суждений (предложений) и как логику вероятностей (исчисление вероятностей)».

В этом пространном документе [25], допущенном в качестве учебника для философских факультетов университетов, всего одним предложением отмечена роль Платона Сергеевича Порецкого такими словами: …«работы казанского математика и астронома П.С. Порецкого (1846-1907) сыграли определенную роль в совершенствовании ряда теоретических и технических аспектов алгебры логики».

Вызывает законное удивление, что за последние 1.5 века ни один крупный математик так и не высказался по вопросу связи математической логики и теории вероятностей, что давала повод различным критикам говорить, что ни в одном из крупных трактатах по теории вероятностей *нет никакого упоминания о существовании какой-либо связи* этой научной дисциплины с алгеброй логики. Имелись ввиду монографии академиков А.А.Маркова, А.Н.Колмогорова и других.

Возникает вопрос: в чем заключается феномен логико-вероятностного анализа и его *замалчивания математиками?* Так в учебнике «Введение в математическую логику» [26] нет даже упоминания о Порецком П.С.

В учебнике [25] не сказано, что П.С.Порецкий являлся самым ярким представителем логической мысли не только Казанского университета, но всей России и мировой науки, что он достиг мировой известности и признания, что его работы существенно развили достижения Буля, Джевонса и Шрёдера.

Дадим высокую оценку путей мышления автора в последней четверти 19 века, которые во второй половине 20 века привели к

важным открытиям в области логико-вероятностного анализа [27, 28]. Здесь уместно вспомнить оценку профессора С.А.Яновской деятельности П.С.Порецкого в масштабной работе «30 лет математики в СССР»…«Независимо от Лейбница идеи алгебры логики, или исчисления классов, равносильного логике Аристотеля, были развиты наряду со многими другими исчислениями, созданными в XIX столетии, А. де Морганом, Булем, Джевонсом, Пирсом, Шредером. Венцом этого периода в истории математической логики были работы русского логика, астронома и математика, собрата Н.И.Лобачевского по Казанскому университету Платона Сергеевича Порецкого» [29].

10 августа 1907 года умер П.С. Порецкий, имя которого более известно за границей, чем на его родине, писал в некрологе профессор И.В. Слешинский – переводчик на русский язык книги Л. Кутюра «Алгебра логики».

Публикация через 130 лет труда П.С.Порецкого «Решение общей задачи теории вероятностей при помощи математической логики» в нашем журнале станет своеобразным интеллектуальным памятником великому русскому логику – первооткрывателю логико-вероятностного анализа Платону Сергеевичу Порецкому.

Литература

1. *Порецкий П.С.* О способах решения логических равенств и об обратном способе математической логики // Собрание протоколов заседаний секции физико-математических наук общества естествоиспытателей при Казанском университете. Казань: 1884. Т. 2. 170 с.
2. *Порецкий П.С.* Изложение основных начал математической логики в возможно более наглядной и общедоступной форме. Сообщение, читанное в 3 заседании секции физико-математических наук общества естествоиспытателей при Казанском университете // Собрание протоколов заседаний секции физико-математических наук общества естествоиспытателей при Казанском университете. Казань: 1881. Т. 1. С. 2–31.
3. *Порецкий П.С.* Решение общей задачи теории вероятностей при помощи математической логики// Собрание протоколов заседаний секции физико-математических наук общества естествоиспытателей при Казанском университете. Казань: 1887. Т.5. С. 83–116.
4. *Boole G.* An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities // London: MacMillan. 1854.
5. *Бажанов В.А.* П.С. Порецкий. Жизнь и научная деятельность пионера исследований в области математической логики в России // Вопросы истории естествознания и техники. 2005. №4. С. 64–73.
6. *Голота Я.Я.* О двух «вычислительных вольностях», огорчающих логика. URL: www.inftach.webservis.ru/it/conference/scm/2000/session 4/golota 2.html. (дата обращения: 01.09.2015).
7. *Соколюк В.Н.* Парадоксы современного бытия (об адекватности логики, мышления мировоззренческим принципам) // Философский век. Альманах. Между физикой и метафизикой: Наука и философия. СПб. 1998. Вып. 7. С.100–107.

8. *Бернштейн С.Н.* Опыт аксиоматического обоснования теории вероятностей // Сообщения Харьковского Математического общества. 1917. Том XV. Сер. 2. С. 209–274.
9. *Мерекин Ю.В.* Решение задач вероятностного расчета однотактных схем методом ортогонализации // Вычислительные системы. Сборник трудов Института СО АН СССР. 1963. Вып.4. С.10–21.
10. *Рябинин И.А.* Основы теории и расчета надежности судовых электроэнергетических систем // Изд-во «Судостроение». Ленинград. 1967. 362 с.
11. *Fratta L., Montanari U.G.* A Boolean Algebra Method for Computing the Terminal Reliability in a Communication Network // IEEE Trans. Circuit Theory. 1973. vol. CT-20. pp. 203–211.
12. *Рябинин И.А., Струков А.В.* Кратко аннотированный список публикаций зарубежных периодических изданий по вопросам оценивания надежности структурно-сложных систем // Труды международной научной школы «Моделирование и анализ безопасности и риска в сложных системах» (МАБР-2011). СПб. 2011. С. 363–379.
13. *Стяжкин Н.И.* Формирование математической логики // М.: «Наука». 1967. 508 с.
14. *Williamson J., Gabbay D.* Editorial. Special issue on Combining Probability and Logic // Journal of Applied Logic. 2003. vol. 1. Issues 3–4. pp. 135–138.
15. *Cozman F., et al.* Special issue on combining probability and logic introduction. 2006. URL: http://www.philos.rug.nl/~romeyn/paper/2009_progicnet_-_editorial_JAL.pdf дата обращения 12.08.15).
16. *Landes J., Williamson J.* Special issue: Combining probability and logic // Journal of Applied Logic. 2015. URL: http://www.sciencedirect.com/science/article/pii/S1570868315000786 (дата обращения 12.08.2015).
17. *Nilsson N.J.* Probabilistic Logic // Artificial Intelligence. Elsevier Science Publ. 1986. vol. 28. pp. 31–56.
18. *Ryabinin I.* Reliability of engineering systems. Principles and Analysis // MIR Publieshers. Moscow. 1976. 531 p.
19. *Рябинин И.А.* Надежность и безопасность структурно-сложных систем // Изд-во С.-Петербурского университета. 2007. 276 с.
20. *Keynes J.M.* Treatise on Probability // L-N.Y.:1921.
21. *Нейман Дж.* Вероятностная логика и синтез надежных организмов из ненадежных компонент // Сб. Автоматы/ М.: ИЛ. 1956. С.68–139.
22. *Сперанский С.О.* Логика вероятности и вероятностная логика // Диссертация к.ф-м.н. Новосибирск. 2013. 109 с.
23. *Рябинин И.А.* О связи математической логики с теорией вероятностей // Ученые записки РГГМУ. СПб. 2008. №6. С.170–176.
24. *Rouche N.* Extension du formalizme f'algebre logique // Revue H.F. 1956. vol. 3. no. 5. p.179–182.
25. Формальная логика: учебник / под ред. Чепухина И.Я., Бродского И.Н. // Л. Издательство Ленинградского университета. 1977. 357 с.
26. *Колмогоров А.Н., Драгалин А.Г.* Введение в математическую логику // М:. Изд. Московского университета. 1982. 120 с.
27. *Рябинин И.А.* Логико-вероятностный анализ проблем надежности и безопасности // Saarbrucken. Academic Publishing. 2012. 263 p.
28. *Ryabinin I.A.* Logical probabilistic analysis and its history // Int. J. of Risk Assessment and Management. 2015. vol.18. no.3/4. pp. 256–265.
29. Математика в СССР за тридцать лет. 1917-1947 / под ред. Курош А.Г., Маркушевич, А.И., Рашевский П.К. // Изд-во ГИТТЛ. 1948. 1044 с.

References

1. Poreckij P.S. [On methods for solving logical equations and the inverse method of mathematical logic]. *Sobranie protokolov zasedaniy sektsii fiziko-matematicheskih nauk obschestva estestvoispyitateley pri Kazanskom universitete* [Collection of Records of Meetings of the Section for Physic and Mathematics of the Scientific Society of the Kasan University]. Kasan. 1884. vol. 2. no. XXIV. 170 p. (In Russ.).
2. Poreckij P.S. [Presentation of fundamental principles of mathematical logic in more evident and popular form] *Sobranie protokolov zasedaniy sektsii fiziko-matematicheskih nauk obschestva estestvoispyitateley pri Kazanskom universitete* [Presentation at the Third meeting of the Kasan Society for Natural Sciences, Collection of Records of Meetings of the Section for Physic and Mathematics of the Scientific Society of the Kasan University] Kasan. 1881. vol. 1. pp. 2–31. (In Russ.).
3. Poreckij P.S [Solving general tasks in Probability Theory by using Mathematical Logic] *Sobranie protokolov zasedaniy sektsii fiziko-matematicheskih nauk obschestva estestvoispyitateley pri Kazanskom universitete* [Collection of Records of Meetings of the Section for Physic and Mathematics of the Scientific Society of the Kasan University] Kasan. 1887. vol. 5 pp. 83–116. (In Russ.).
4. Boole G. An investigation of the laws of thought, on which are founded the mathematical theories of logic and probabilities. London: MacMillan. 1854.
5. Bazhanov V.A. [P.S.Poreckij. Life and work of pioneer of Mathematical Logic studies in Russia] *Voprosy Istorii Estestvoznaniya i tekhniki – Questions of History of Science and Technology*. 2005. vol. 4. pp. 64–73. (In Russ.).
6. Golota Ja.Ja. O dvuh «vychislitel'nyh vol'nostjah», ogorchajushhih logika [About two "computational liberties," grieve logic] Available at: http://www.inftech.webservis.ru/it/ conference/ scm /2000/session 4/golota 2.html. (accessed 1.09.2015). (In Russ.).
7. Sokoljuk V.N. [The paradoxes of modern life (the adequacy of logic, thinking philosophical principles)] *Philosophskii vek. Almanah. Mezshdu phizikoi I metaphizikoi: Nauka I philosophia – The Philosophical Age. Almanac. Between Physics and Metaphysics: Science and Philosophy.* SPb. 1998. vol. 7. pp. 100–107. (In Russ.).
8. Berstien S.N. [Experience axiomatic foundation of the theory of probability] *Soobscheniya Harkovskogo Matematicheskogo obshcestva – Communications of the Kharkov Mathematical Society*. 1917. vol. XV. Issue 2. pp. 209–274. (In Russ.).
9. Merekin Ju.V. [Problem solving probabilistic calculation of single-ended circuits by orthogonalization] *Vychislitelnye sistemy. Sbornik trudov Instituta SO AN SSSR – Computer systems. Proceedings of the Institute of SB RAS.* 1963. vol. 4. pp. 10–21. (In Russ.).
10. Ryabinin I.A. *Osnovy teorii i rascheta nadezhnosti sudovyh jelectro-energeticheskih system* [Fundamentals of the theory and calculation of reliability of ship electric power systems]. Izd-vo «Sudostroenie». Leningrad. 1967. 362 p. (In Russ.).
11. Fratta L., Montanari U.G. A Boolean Algebra Method for Computing the Terminal Reliability in a Communication Network. IEEE Trans. Circuit Theory. 1973. vol. CT-20. pp 203–211.
12. Ryabinin I.A., Strukov A.V. [Briefly annotated list of publications of foreign periodicals concerning estimation of reliability of structural complex systems] *Trudy mezhdunarodnoy nauchnoy shkoly «Modelirovanie i analiz bezopasnosti i riska v slozhnyih sistemah» (MABR-2011)* [Proceedings of the international school of sciences "Modeling and the analysis of safety and risk in complex systems" (MABR-2011)] 2011.SPb. pp. 363–379. (In Russ.).

13. Styazhkin N.I. *Formirovanie matematicheskoj logiki* [Formation of the mathematical logic]. M.: Nauka. 1967. 508 p. (In Russ.).
14. Williamson J., Gabbay D. Editorial. Special issue on Combining Probability and Logic. *Journal of Applied Logic*. 2003. vol. 1. Issues 3–4. pp. 135–138.
15. Cozman F., et al. Special issue on combining probability and logic introduction. 2006. Available at: http://www.philos.rug.nl/~romeyn/paper/2009_progicnet_-_editorial_JAL.pdf (accessed 12.08.15).
16. Landes J., Williamson J. Special issue: Combining probability and logic // Journal of Applied Logic. 2015. Available at: http://www.sciencedirect.com/science/article/pii/S1570868315000786 (accessed 12.08.2015)
17. Nilsson N.J. Probabilistic Logic. *Artificial Intelligence*. Elsevier Science Publ. 1986. vol. 28. pp. 31–56.
18. Ryabinin I. Reliability of engineering systems. Principles and Analysis. M. MIR Publieshers. 1976. 531 p.
19. Ryabinin I.A. *Nadeznost' i bezopasnost' strukturno-slozhnyh sistem* [Reliability and safety of structural complex systems.]. SPb.: S.-Peterb. Un-ta, 2007. 276 p. (In Russ.).
20. Keynes J.M. Treatise on Probability. L-N.Y.: 1921.
21. Neumann J. [Probabilistic logic and synthesis of reliable organisms from unreliable component]. Sb.Automaty/ M. : IL. 1956. pp.68–139. (In Russ.).
22. Speranskii S.O. *Logika verojatnosti i verojatnostnaja logika* [The logic of probability and probabilistic logic]. Ph.D. thesis phis-math. science. Novosibirsk. 2013. 109p. (In Russ.).
23. Ryabinin I.A. *O svjazi matematicheskoj logiki s teoriej verojatnostej* [On the relationship between mathematical logic, probability theory]. *Uchenye zapiski PGGMU – Proceedings of the RSHU. A theoretical research journal*. SPb. 2008. vol. 6. pp. 170–176. (In Russ.).
24. Rouche N. Extension du formalizme f'algebre logique. *Revue H.F*. 1956. vol. 3. no. 5. pp. 179–182. (In France).
25. *Formal'naja logika: uchebnik. Pod red. Chepuhina I.Ja., Brodskogo I.N.* [Formal logic: textbook. Edited by Chepuhin I.Ja., Brodskij I.N.]. L. Izdatelstvo Leningradskogo Universiteta. 1977. 357 p. (In Russ.).
26. Kolmogorov A.N., Dragalin A.G. *Vvedenie v matematicheskuju logiku* [Introduction to mathematical logic]. M:. Izd. Moskovskogo Universiteta. 1982. 120 p. (In Russ.).
27. Ryabinin I.A. *Logiko-verojatnostnyj analiz problem nadezhnosti i bezopasnosti* [Logical and probabilistic analysis of the problems of reliability and safety]. Saarbrucken. Academic Publishing. 2012. 263 p. (In Russ.).
28. Ryabinin I.A. Logical probabilistic analysis and its history. *Int. J. of Risk Assessment and Management*. 2015. vol. 18. no. 3/4. pp. 256–265.
29. *Matematika v SSSR za tridcat' let. 1917-1947. Pod red. Kurosh A.G., Markushevich A.I., Rashevskij P.K.* [Mathematics in the USSR over thirty years. 1917-1947. Edited by Kurosh A.G., Markushevich A.I., Rashevskii P.K]. Izd-vo GITTL. 1948. 1044 p. (In Russ.).

Рябинин Игорь Алексеевич — д-р техн. наук, профессор, почетный профессор Военно-морской академии, профессор, Военно-морская академия им. Н.Г. Кузнецова. Область научных интересов: анализ данных, системный анализ, теория надежности, модели и методы анализа надежности и безопасности структурно-сложных технических систем. Число научных публикаций — 229. Ryabinin25@mail.ru; 26-я линия В.О., дом 15, корп.2, Санкт-Петербург, 199155; р.т.: +7(812)5556570.

Ryabinin Igor Alekseevich — Ph.D., Dr. Sci., professor, Professor Emeritus, professor, N.G. Kuznetsov Naval Academy. Research interests: data analysis, analysis of systems,

reliability theory, mathematical models and methods reliability and safety analysis of structurally-complex systems. The number of publications — 229. Ryabinin25@mail.ru; 15/2, 26th line of Vasilievsky Island, St. Petersburg, 199026; office phone: +7(812)5556570.

Струков Александр Владимирович — к-т техн. наук, доцент, ведущий инженер исследовательского отдела, АО Специализированная инжиниринговая компания «Севзапмонтажавтоматика» (СПИК СЗМА). Область научных интересов: анализ данных, системный анализ, теория надежности, модели и методы принятия решения в сложных организационно- технических системах. Число научных публикаций — 71. alexander_strukov@szma.com; 26-я линия В.О., дом 15, корп.2, Санкт-Петербург, 199155; р.т.: +7(812) 610 78 74, Факс: +7(812) 610 78 74.

Strukov Alexandr Vladimirovich — Ph.D., associate professor, senior engineer of research department, public corporation Specialized engineering company "Sevzapmontageautomatica" (SPIK SZMA). Research interests: data analysis, analysis of systems, reliability theory, mathematical models and methods of decision-making support in complex technical-organizational systems. The number of publications — 71. alexander_strukov@szma.com; 15/2, 26th line of Vasilievsky Island, St. Petersburg, 199026; office phone: +7(812) 610 78 74, Fax: +7(812) 610 78 74.

РЕФЕРАТ

Рябинин И.А., Струков А.В. **Предисловие и вступительная статья к переизданию работы П.С. Порецкого «Решение общей задачи теории вероятностей при помощи математической логики».**

Предисловие и вступительная статья к переизданию работы П.С. Порецкого «Решение общей задачи теории вероятностей при помощи математической логики» имеют целью представить работу выдающегося русского математика, астронома и логика, изданную в 1887 г. в Императорском Казанском Университете. Высказывая глубокое уважение трудам Дж. Буля, которые заложили основы новой отрасли знаний, П.С. Порецкий дополняет и развивает его идеи и идеи его учеников о возможности применения математической логики к другим областям науки и к теории вероятностей в частности. Уже в самом начале работы на философский вопрос: возможно ли приложение учения о качественных символах (логических классах) к учению о символах количественных (вероятностных)? И отвечает: возможно. Опираясь на теоремы о произведении независимых событий и сумме несовместных событий, П.С. Порецкий описывает процедуру *перехода* от логического равенства между событиями к алгебраическому равенству между их вероятностями, давая тем самым первое определение логико-вероятностного анализа. В современной литературе переход к дисъюнктивному (по выражению П.С. Порецкого) виду логической формулы получил название метода ортогонализации или метода Sum of Disjoint Products (сумма несовместных произведений). Потребности практики привели к необходимости «повторного открытия» этого метода в 1963 г. Ю.В. Мерекиным для решения задач анализа однотактных цифровых схем. Затем в 1967г. в СССР (И.А. Рябинин) и в 1973г. итальянские ученые L. Fratta и U. Montanari начали применять этот метод для анализа надежности структурно-сложных систем, тем самым подтверждая мысль физика П.Эренфеста о перспективности применения математической логики в технике. Приведены примеры вычисления вероятностей на сложных структурах. Выяснение взаимоотношений между вероятностью и логикой, поиск различий и общего в понятиях *вероятностная логика* и *логика вероятностей* остается актуальным и сейчас, регулярно обсуждается, дискутируется на семинарах, конференциях, в научной литературе. Переиздание работы П.С. Порецкого может стать важным историческим вкладом в область вероятностного анализа при помощи логических методов.

SUMMARY

Ryabinin I.A., Strukov A.V. **A Preface and an Introductory Article to the Re-edition of the Work of Platon Sergeevich Poreckiy «Solving General Tasks in Probability Theory by Using Mathematical Logic».**

A preface and an introductory article to the re-edition of the work of P.S. Poreckiy, published in 1887 in the Imperial University of Kazan, “Solving general tasks in probability theory by using mathematical logic” are aimed at presenting the work of the eminent Russian mathematician, astronomer and logician. Expressing deep respect to George Boole’s works which laid the foundations for a new branch of knowledge, P.S. Poreckiy supplements and develops his and his students’ ideas that it’s possible to apply the mathematical logic to other areas of science and probability theory in particular. At the very beginning of his work there is a philosophical question - is the application of the theory of quality symbols (logic classes) to the theory of quantitative (probability) symbols possible? And his answer to it is: it is possible. Based on the theorems on the product of independent events and the amount of incompatible events, P.S. Poreckiy describes the procedure of transition from the logic equation between the events to algebraic equality between their probabilities, thus giving the first definition of logic-and-probabilistic analysis. In modern literature, the transition to disjunctive (in the words of P.S. Poreckiy) view of logical form is called orthogonalization method or the method of Sum of Disjoint Products (amount of incompatible products). Requirements of practice have led to the necessity to "re-open" this method in 1963 by Y.V. Merekin in order to solve the problems of single-cycle digital circuits analysis. Then, in 1967 in the USSR I.A. Ryabinin and in 1973 Italian scientists L. Fratta and U. Montanari started to apply this method for reliability analysis of structurally complex systems, thus confirming the idea of physicist P.Erenfest about the prospective application of mathematical logic in technique. The examples of probability calculations on complex structures are given. Clarification of the relationship between probability and logic, search of differences and similarities in terms of probability logic and the logic of probability remains relevant today and it is regularly discussed, debated at seminars, conferences and in the scientific literature. P.S. Poreckiy’s work re-editing can be an important historical contribution to the field of probabilistic analysis using the logic methods.

УДК 621.382

П.С. Порецкий

РЕШЕНИЕ ОБЩЕЙ ЗАДАЧИ ТЕОРИИ ВЕРОЯТНОСТЕЙ ПРИ ПОМОЩИ МАТЕМАТИЧЕСКОЙ ЛОГИКИ

Порецкий П.С. **Решение общей задачи теории вероятностей при помощи математической логики.**
Аннотация. Сообщение П.С.Порецкого, читанное 25 октября 1886г. на 60-м заседании секции физико-математических наук Общества Естествоиспытателей при Императорском Казанском Университете. Печатается в авторской редакции 1886 года (Порецкий П.С. Решение общей задачи теории вероятностей при помощи математической логики. - Собрание протоколов 60-го заседания секции физико-математических наук общества естествоиспытателей при Казанском университете, Казань, 1886, С. 1-34.).
Ключевые слова: теория вероятностей, математическая логика.

Poreckii P.S. **Solving General Tasks in Probability Theory by Using Mathematical Logic.**
Abstract. Lecture of P.S. Poreckii hold on October, 25th 1886, at the 60th Meeting of Section for Physic and Mathematics of the Scientific Society of the Imperial Kazan University. It is published in the original edition of 1886 (Poretsky P.S. Solution of the general problems of probability theory with the help of mathematical logic. - The meeting protocols of the 60th meeting of the Section of Physics and Mathematics Society of Naturalists at Kazan University, Kazan, 1886, pp 1-34).
Keywords: theory of probabilities; mathematical logic.

В 1884 году я публиковал сочинение «О способах решения логических равенств», где изложена полная теория этих равенств.

Здесь я предполагаю применить эту теорию к решению следующей задачи Теории Вероятностей: определить вероятность сложного события, зависящего от данных простых событий, с помощью вероятностей всех или нескольких (произвольно избранных) из этих простых событий, а также вероятностей некоторых других сложных событий, предполагая, что данные события подчинены произвольному числу каких бы то ни было условий.

Очевидно, это есть самая общая задача относительно определения вероятностей событий. Сколько мне известно, в Теории Вероятностей нет способа решения этой задачи в

общем виде. А потому решение ее с помощью Математической Логики не должно представляться излишним.

Решение этой задачи, данной Булем в его сочинении An investigation of the laws of thought, нельзя считать научным, как потому что оно основано на произвольной и чисто эмпирической теории логических равенств, так и потому, что самая идея о переходе от логических равенств к алгебраическим разработана у Буля неудачно. Таким образом, главная цель настоящей статьи – дать научную форму глубоко, но смутной и бездоказательной, идеи Буля о применимости Математич. Логики к Теории Вероятностей.

§1. Прежде всего возникает вопрос: возможно ли приложение учения о качественных символах (логических классах) к учению о символах количественных (вероятностях)? Отвечаем: возможно.

В самом деле, логическое равенство

$$f(a,b,c,\partial,...) = \varphi(a,b,c,\partial,...)$$

означает, что в пределах некоторого мира речи, все предметы, относящиеся к классу f, вполне тождественны с предметами класса φ, и что все отличие между классами f и φ заключается в различной классификации одних и тех же предметов. Если так, то *число* предметов, содержащихся в классах f и φ, должно быть одно и то же, т. е. напр.

$$N[f(a,b,c,\partial,...)] = N[\varphi(a,b,c,\partial,...)].$$

Вот чисто математическое равенство, прямо вытекающее из исходного логического. Отсюда уже легко перейти и к отношению между вероятностями. Если означим через $N(1)$ число всех предметов мира речи и назовем отношение $N(f)/N(1)$, т.е. вероятность класса f, символом $P(f)$, то понятно, что

$$P[f(a,b,c,\partial,...)] = P[\varphi(a,b,c,\partial,...)].$$

И так, если два класса логически равнозначны, то их вероятности равны между собою.

Отсюда открывается следующий общий путь для определения вероятностей: найти логическую связь между событием, которого вероятность ищется, и другими событиями, вероятности которых даны, а затем сделать *переход* от логического равенства между событиями к алгебраическому равенству между их вероятностями.

Построением правил для такого перехода от логического равенства к соответственному алгебраическому нам и предстоит теперь заняться.

§2. Пусть логические символы $a,b,c,...$ означают простые события. В таком случае, логические отрицания тех же символов, т.е.

$a_o, b_o, c_o, \ldots$, должны означать соответственно: всякое, в пределах мира речи, событие, только не a ; всякое событие, кроме b , и т.д. Затем, логические суммы в роде $a+b, a+b_o$ и т.д. должны означать сложные события, состоящие: первое - в наступлении или a , или b ; второе – в наступлении или a , или всякого события, кроме b , и т.д. Наконец, логические произведения вроде ab, ab_o и т.д. должны означать сложные события, состоящие: первое – в совпадении событий a и b , второе – в совпадение события a с каким угодно событием, кроме b , и т.д.

Понятно, что например, логическое выражение

$$a+b(c_o+\partial_o)+b_o\partial_o$$

означает сложное событие, которое наступает: во 1-х при наступлении события a ; во 2-х, при совпадении события b или с событием не-c , или же с событием не-∂ ; и наконец, в 3-х, при совпадении события не-b с событием не-∂ .

§3. Из Теории вероятностей известно, что вероятность ненаступления события равна единице (достоверности) без вероятности его наступления.

Если так, то

$$P(a_o)=1-P(a).$$

Точно также, например,

$$P(b_o)=1-P(b).$$

и пр.

§4. Далее, из Теории Вероятностей известно, что если два события несовместны, то вероятность, что случится то или другое из них, равна сумме их отдельных вероятностей. Поэтому, если логические классы m и n дисъюнкты, т.е. не имеют общих предметов, (причем $mn=0$), то

$$P(m+n)=P(m)+P(n).$$

Это правило применимо к какому угодно числу несовместных одно с другим событий. Для возможности пользоваться этим правилом необходимо уметь каждый логический многочлен

$$A+B+C+D+\ldots$$

приводить к дисъюнктному виду, т.е. к виду

$$A+A_oB+A_oB_oC+A_oB_oC_oD+\ldots,$$

где A_o есть отрицание A , B_o – отрицание B и т.д. Оба написанные многочлена логически равнозначны, но отличаются тем, что к первому из них не применимо предыдущее правило, тогда как во второму применимо.

И так, каждое сложное событие, имеющее вид суммы, мы всегда можем выразить так, что его вероятность разобьется на сумму вероятностей других, более простых событий. Напр., вероятность

$$P(A+B+C+D),$$

будучи приведена к виду

$$P(A+A_oB+A_oB_oC+A_oB_oC_oD),$$

разбивается на сумму вероятностей:

$$P(A)+P(A_oB)+P(A_oB_oC)+P(A_oB_oC_oD).$$

§5. Затем, из Теории Вероятностей известно, что если два и более события суть независимы, то вероятность их совпадения равна произведению их отдельных вероятностей. Это означает, что если $a,b,c,...$ суть простые события, не связанные между собою никаким логическим отношением, то

$$P(abc...)=P(a)P(b)P(c)...$$

§6. Если так, то вероятность приведенного к дисъюнктному виду логического многочлена

$$A+A_oB+A_oB_oC+...,$$

не подчиненного никаким условиям, может быть изображена так:

$$P(A)+P(A_o)P(B)+P(A_o)P(B_o)P(C)+...,$$

т.е. получается из выражения многочлена простою заменой классов $A,B,C,...$ и их отрицаний вероятностями тех и других.

Отсюда видим, что абсолютная вероятность всякой отдельной логической функции

$$f(a,b,c,\partial,...),$$

приведенной предварительно к дисъюнктному виду, есть

$$f[P(a),P(b),P(c),...].$$

В первом из этих выражений f означает совокупность *логических* действий над качеств. символами $a,b,c,...$; во втором то же f означает совокупность *алгебраических* действий над количеств. символами $P(a),P(b),P(c),...$.

Пример. Если вероятности простых событий x и y суть $P(x)=p$, $P(y)=q$, то вероятность сложного события xy_o+x_oy, уже имеющего дисъюнктный вид, есть $p(1-q)+(1-p)q$. Вероятность же сложного события $x+y$, которое, по приведении к дисъюнкт-

ному виду, есть $x + x_o y$ или $y + y_o x$ выразится так: $p + (1-p)q$, или $q + (1-q)p$.

Так делается переход от выражения отдельной логической функции к выражению абсолютной её вероятности.

§ 7. Понятно теперь, что для перехода от логического равенства $f = \varphi$ к отношению между вероятностями входящих туда классов, надо привести обе функции f и φ к дисъюнкткому виду и затем заменить в обеих частях равенства все качественные символы $a, b, c, \ldots$ символами количественными $P(a), P(b), \ldots$

Для примера превратим логическое равенство

$$ab + c\partial = ac = b\partial$$

в отношение между вероятностями, принимая

$$P(a) = p,\ P(b) = q,\ P(c) = r,\ P(\partial) = s\ .$$

Надо привести к дисъюнктному виду обе части исходного равенства. Имеем:

$$ab + (ab)_o c\partial = ac + (ac)_o b\partial;$$
$$ab + (a_o + b_o)c\partial = ac + (a_o + c_o)b\partial;$$
$$ab + (a_o + ab_o)c\partial = ac + (a_o + ac_o)bd;$$
$$ab + a_o c\partial + ab_o c\partial = ac + a_o b\partial + ac_o bd.$$

В последнем равенстве обе части состоят из членов, дисъюнктных между собою, а потому, делая от него переход к отношению между вероятностями, получим:

$$pq + (1-p)rs + p(1-q)rs = pr + (1-p)qs + p(1-r)qs\ .$$

§ 8. Хотя, таким образом, при операциях над логическими равенствами мы можем в любой момент сделать переход к отношениям между вероятностями; однако, при решении задачи об определении вероятности одного события посредством вероятностей других событий, всего натуральнее поступать так: найти из всей совокупности данных логич. условий определение первого события с помощью остальных и уже затем сделать переход к вероятностям. Этого приема мы и будем держаться.

§ 9. Доселе мы вели речь об абсолютных вероятностях. Обращаемся к вероятностям относительным.

В Теории Вероятностей доказывается следующая истина: вероятность, что если событие ***A*** случится, то и событие ***B*** тоже случится,

равна вероятности совпадения событий ***A*** и ***B***, разделенной на вероятность события ***A***, т. е. равна дроби $\frac{P(AB)}{P(A)}$.

Поэтому, если $A = f(a,b,c,\partial,...)$, $B = \varphi(a,b,c,\partial,...)$, то искомая относительная вероятность получится, если в выражение произведения f и φ, приведенного к дисъюнктному виду, заменим $a,b,c,...$ их абсолютными вероятностями и полученный результат разделим на выражение функции f, приведенное к дисъюнктному виду, причем в нем надо также заменить все качественные символы количественными.

И так, искомая относит. вероятность будет:

$$\frac{[f(a,b,c,...)\varphi(a,b,c,\partial...)]}{[f(a,b,c,\partial...)]},$$

где заключение в прямые скобки означает упомянутую замену.

Для примера, полагая $P(x) = p, P(y) = q, P(z) = r$, найдем вероятность, что если случится событие

$$xy_o + x_o y,$$

т.е. одно из событий x и y, но не оба вместе, то случится также и событие

$$yz_o + y_o z,$$

т.е. одно из событий y и z, но не оба вместе.

В данном случае

$$f(x,y,z) = xy_o + x_o y, \quad \varphi(x,y,z) = yz_o + y_o z$$
$$f(x,y,z)\varphi(x,y,z) = xy_o z + x_o yz_o .$$

След. искомая относит. вероятность есть:

$$\frac{[f\varphi]}{[f]} = \frac{p(1-q)r + (1-p)q(1-r)}{p(1-q) + q(1-p)}.$$

§ 10. Предположим теперь, что даны относительные вероятности простых событий $a,b,c,...$, высчитанные так, чтобы удовлетворялся ряд условий

$$f'(a,b,c,...) = \varphi'(a,b,c,...), \; f'' = \varphi'', \; f''' = \varphi''', ...,$$

и требуется найти абсолютные вероятности тех же простых событий.

Заметим, прежде всего, что каждое логическое равенство

$$f(a,b,c,...) = \varphi(a,b,c,...)$$

может быть тождественно заменено равенством

$$1 = f\varphi + f_o\varphi_o,$$

где 1 означает логический мир речи (в данном случае, мир всех событий, о которых идет речь), f_o и φ_0, суть отрицания f и φ.

Кроме того, известно, что вся совокупность данных условий вполне равнозначна с одним условием:

$$1 = (f'\varphi' + f'_o\varphi'_o)(f''\varphi'' + f''_o\varphi''_o)(f'''\varphi''' + f'''_o\varphi'''_o)\ldots,$$

которое можно сокращенно представить под формой

$$1 = M(a,b,c,\partial\ldots).$$

В этом равенстве, тождественно заменяющем все данные условия, функция M называется логическим миром речи задачи или полной логич. единицей задачи.

И так, подчинение классов $a,b,c,\ldots$ всей совокупности исходных условий вполне равнозначно с подчинением их одному условию $1 = M(a,b,c,\partial\ldots)$, составленному по правилу указанному выше.

Пусть теперь $p,q,r,\ldots$ означают вероятности событий $a,b,c,\ldots$, подчиненных условию $1=M$, и пусть $p',q',r',\ldots$ абсолютные вероятности тех же событий. Так как первые из этих вероятностей означают вероятности, что при наступлении события M (мира речи) случатся события $a,b,c,\ldots$, то, по доказанному ранее, для определения абсолютных вероятностей $p',q',r',\ldots$, будем иметь:

$$p = \frac{[aM]}{[M]},\ q = \frac{[bM]}{[M]},\ r = \frac{[cM]}{[M]},\ldots,$$

где в правых частях классы $a,b,c,\ldots$ должны быть заменены их абсолютными вероятностями $p',q',r',\ldots$, которые и найдутся чрез решение системы полученных алгебраич. уравнений.

Возьмем примерь. Пусть при вынимании из урны шаров обращали внимание только на случаи, когда вынутый шар был или белый, или мраморный (или то и другое вместе), и пусть, при этом условии, найдены: p -вероятность белого шара, q - мраморного. Найти абсолютные их вероятности p' и q'.

Построим сначала условие, с подчинением которому были найдены вероятности p и q. Пусть x есть вынутие белого шара, y - мраморного. Если при высчитывании вероятностей исключались случаи, когда вынутый шар был не белый и не мраморный, то это значит, что было соблюдаемо условие:

$$x_o y_o = 0,$$

или, что тоже:

$$1 = xy + x_o y + x y_o .$$

И так, в данном случае

$$M(x,y) = xy + x_o y + x y_o$$
$$xM(x,y) = xy + x y_o = x$$
$$yM(x,y) = xy + x_o y = y \ .$$

А потому имеем:

$$p = \frac{[Mx]_{x=p',y=q'}}{[M]_{x=p',y=q'}} \quad , q = \frac{[My]_{x=p',y=q'}}{[M]_{x=p',y=q'}},$$

или:

$$p = \frac{p'}{p'q' + p'(1-q') + q'(1-p')}, \ q = \frac{q'}{p'q' + p'(1-q') + q'(1-p')}.$$

Через решение этих двух алгебр. уравнений получим

$$p' = \frac{p+q-1}{q}, \ q' = \frac{p+q-1}{p}.$$

§ 11. Согласно с тем, что высказано ранее, для определения вероятности одного события через вероятности других событий, нам надо прежде всего логически выразить первое через остальные. Это нас заставляет сказать несколько слов о приемах определения одного логического класса (простого или сложного) через все или некоторые из прочих.

Пусть требуется определить простой класс a через все прочие классы $b, c, \partial, \ldots$, связанные с a и между собою рядом условий (посылок):

$$f' = \varphi', \ f'' = \varphi'', \ f''' = \varphi''', \ldots .$$

Все эти условия тождественно могут быть заменены одним:

$$1 = M(a,b,c,\partial,..) .$$

С другой стороны, это последнее равенство может быть тождественно замещено следующими тремя:

$$a = aM(1,b,c,\partial...) = aM(1)$$
$$a = a + M(1,b,c,,...)M_o(0,b,c,...) = a + M(1)M_o(0) \cdot$$
$$1 = M(1,b,c,,...) + M_o(0,b,c,...) = M(1) + M_o(0).$$

Здесь $M(1)$ есть результат замещения в функции $M(a,b,c,...)$ класса a единицей, а его отрицания a_o нулем, $M(0)$ есть результат за-

мещения в $M(a,b,c,...)$ класса *a* нулем, а его отрицания a_o единицей; $M_o(0)$ есть отрицание функции $M(0)$ или, что тоже, результат замещения в отрицании функции *M,* т. е. в функции $M_o(a,b,c,...)$, класса *a* нулем, а его отрицания a_o единицей.

Из последних трех равенств первое показывает, что *a* содержится в $M(1)$, второе - что *a* содержит в себе $M_o(0)M(1)$. Вот почему эти два равенства можно заменить неравенствами

$$a < M(1),\ a > M_o(0)M(1),$$

которые надо понимать в смысле: *a* не больше $M(1)$ и не меньше $M_o(0)M(1)$.

Наконец, третье равенство $1 = M(1) + M(0)$, зависящее от классов $b,c,\partial,...$, но не содержащее класса *a*, представляет условие, которому, в силу первоначальных условий, подчинены те две функции $M(1)$ и $M_o(0)M(1)$, с помощью которых определяется *a*.

В случае, когда эти две функции логически равнозначны, т. е. когда

$$M_o(0)M(1) = M(1),$$

два неравенства, определяющие *a*, суть:

$$a > M(1),\ a < M(1).$$

т.е. доставляют одно равенство:

$$a = M(1).$$

Если желаем определить *a* из того же уравнения $1 = M(a,b,c,...)$ не через все, но через некоторые из классов $b,c,\partial,...$, то все лишние классы надо исключить из равенства $1 = M(a,b,c,\partial,...)$. Для этого исключения достаточно заменить в равенстве $1 = M(a,b,c,...)$ все исключаемые классы, а также их отрицания, единицами. Пусть результат исключения будет: $1 = M'$, где M' зависит от *a* и некоторых из прочих классов. Затем останется определить *a* из равенства $1 = M'$ совершенно так, как мы выше определяли его из равенства $1 = M$.

Так определяется простой класс через все или некоторые прочие простые классы на основании какого бы то ни было числа данных логических условий.

§ 12. Обращаемся к определению сложных классов, т. е. функций.

Легко показать, что логическая функция может быть выражена через простые классы (все или некоторые) даже тогда, когда эти последние не подчинены никаким условным равенствам.

В самом деле, пусть даны *n* простые классы $a, b, c, \ldots$, не связанные между собою никакими условиями, и сложный класс ***A***, где ***A*** означает определенную функцию тех же классов. В таком случае, положив $A = w$, или, что тоже, $1 = Aw + A_o w_o$, можем сказать, что мы имеем *n + 1* простых классов: *w,a,b,c,…*, которые подчинены условию

$$1 = Aw + A_o w_o = M(w, a, b, c, \ldots).$$

Из этого условия и может быть определен простой класс *w* (т. е. функция **A**) через все или некоторые из прочих классов по правилам, указанным выше.

Таким образом, рассматривание хотя бы только одной логической функции совместно с независимыми простыми классами обращает задачу из безусловной в условную.

Если, рядом с *n* независимыми простыми классами $w, a, b, c, \ldots$, мы начнем рассматривать *m* функций $U, V, W, \ldots$, то, введя ряд обозначений

$$U = u, V = v, \quad W = w, \ldots,$$

мы получаем задачу об $n + m$ простых классах: $a, b, c, \ldots, u, v, w, \ldots$, подчиненных условию:

$$1 = (uU + u_o U_o)(vV + v_o V_o)(wW + w_o W_o)\ldots = M(a, b, c, \ldots, u, v, w, \ldots),$$

из которого по предыдущему и может быть логически определен любой из классов $u, v, w, \ldots$ с помощью всех или некоторых из прочих классов, т. е. найдется любая из функций $U, V, W, \ldots$ с помощью всех или некоторых из данных простых классов и всех или некоторых из прочих функций.

Наконец, если простые *n* классы $a, b, c, \partial, \ldots$ суть зависимые, связанные между собою *p* условиями

$$A' = B', \quad A'' = B'', \quad A''' = C''', \ldots,$$

где $A', B', A'', B'', \ldots$ суть функции $a, b, c, \partial, \ldots$, то при определении одной из ряда *m* функций

$$U, V, W, \ldots$$

мы будем иметь задачу об $n + m$ простых классах: $a, b, c, \partial, \ldots, u, v, w, \ldots$, связанных между собою $p + m$ условиями

$$A' = B', \quad A'' = B'', \ldots, u = U, \quad v = V, \quad w = W, \ldots,$$

или, что тоже, одним условием:

$$1=(A'B'+A'_o B'_o)(A''B''+A''_o B''_o)\dots(uU+u_oU_o)(vV+v_0V_o)\dots,$$

которое можно изобразить так:

$$1=M(a,b,c,\partial,\dots,u,v,w,\dots).$$

Отсюда и может быть найдена по предыдущему любая из функций $U,V,W,\dots$ с помощью всех или некоторых из прочих функций, а также всех или некоторых из простых классов $a,b,c,\partial,\dots$, причем все исходные условные равенства будут приняты во внимание.

§ 13. Вот мы имеем все данные для решения поставленной в начале статьи общей задачи об определении вероятности одной функции (одного сложного события) посредством вероятностей всех или некоторых прочих функций и простых классов, предполагая, что последние связаны между собою каким бы то ни было числом условных равенств.

Пусть, поступая по предыдущему, мы пришли к равенству

$$1=M(a,b,c,\dots,u,v,w,\dots),$$

из которого уже исключены все классы и функции, вероятности которых не должны быть принимаемы во внимание при определении вероятности функции U с помощью вероятностей прочих классов $a,b,c,\dots,v,w,\dots$

В таком случае мы получим:

$$u<M(1),u>M_o(0)M(1),$$

где $M(1)$ и $M(0)$ суть результаты замещения в функции M класса u единицей и нулем соответственно, (а его отрицания нулем и единицей), причем между прочими классами $a,b,c,\dots,v,w,\dots$ устанавливается отношение:

$$1=M(1)+M_o=K.$$

Остается перейти к определению вероятности u. Пусть вероятности классов $a,b,c,\dots,v,w,\dots$, найденные с соблюдением всех первоначальных условий задачи, а следовательно также подчиненные и условию $1=K$, суть $p,q,r,\dots\alpha,\beta,\dots$ В таком случае их абсолютные вероятности, которые мы назовем через $p',q',r',\dots\alpha',\beta',\dots$, надо искать из условий:

$$p=\frac{[aK]}{[K]},\ q=\frac{[bK]}{[K]},\dots,\alpha=\frac{[vK]}{[K]},\beta=\frac{[wK]}{[K]},\dots,$$

где в правых частях, по приведению числителей и знаменателей к дисъюнктному виду, все качественные символы $a,b,c,\dots,v,w,\dots$ должны быть заменены количественными символами $p',q',r',\dots\alpha',\beta',\dots$

Найденные отсюда величины $p', q', r', ...\alpha', \beta', ..$, будучи подставлены, вместо $a, b, c, ..., v, w, ...$ в правые части неравенств

$$u < M(1), u > M_o(0)M(1),$$

доставят нам абсолютные вероятности функций $M(1)$ и $M(0)M(1)$, т.е. пределы для абсолютной вероятности функции u.

Однако, нам нужно знать не абсолютную, но относительную вероятность функции u, а именно такую, в которой были бы приняты во внимание все условные равенства задачи, а след. также и условие $1 = K$. В силу доказанного ранее, такого рода относительные вероятности функций $M(1)$ и $M(0)M(1)$ суть соответственно:

$$\frac{[M(1)K]}{[K]}, \frac{[M_0(0)M(1)K]}{[K]},$$

где все качественные символы $a, b, c, ..., v, w, ...$ должны быть заменены соответственными абсолютными вероятностями $p', q', r', ...\alpha', \beta', ...$. Но

$$K = M(1) + M(0),$$

а потому

$$M(1)K = M(1)[M(1) + M(0)] = M(1)$$
$$M_o(0)M(1)K = M_o(0)M(1)[M(1) + M(0)] = M_o(0)M(1).$$

Следовательно, относительные вероятности функций $M(1)$ и $M_o(0)M(1)$ суть

$$\frac{[M(1)K]}{[K]} \text{ и } \frac{[M_0(0)M(1)K]}{[K]}.$$

Если так, то, называя искомую относительную вероятность функции u через $P(u)$, получим

$$P(u) < \frac{[M(1)]}{[K]}, \quad P(u) > \frac{[M_0(0)M(1)]}{[K]}, \tag{1}$$

где все качественные символы $a, b, c, ..., v, w, ...$ должны быть заменены символами $p', q', r', ...\alpha', \beta', ...$. После такой замены вместо этих последних символов должны быть подставлены их значения, выраженные с помощью $p, q, r, ...\alpha, \beta, ..$ на основании равенств:

$$p = \frac{[aK]}{[K]}, \quad q = \frac{[bK]}{[K]}, ..., \alpha = \frac{[vK]}{[K]}, \beta = \frac{[wK]}{[K]}, ..., \tag{2}$$

в которых предварительно должно быть сделано то же замещение символов $a, b, c, ..., v, w, ...$ символами $p', q', r', ...\alpha', \beta', ...$. Но если в формулах (1) и (2) качественные символы $a, b, c, ..., v, w, ...$ заменяются количест-

венными символами $p', q', r', \ldots \alpha', \beta', \ldots$, которые вслед затем исключаются из (1) с помощью (2), то понятно, что нет надобности делать означенное замещение на самом деле, а совершенно достаточно начать считать в (1) и (2) качественные символы $a, b, c, \ldots, v, w, \ldots$ как бы количественными и исключить их, по правилам Алгебры, из (1) с помощью (2). Таким образом, окончательная форма решения задачи об определении $P(u)$ с помощью относит. вероятностей $p, q, r, \ldots \alpha, \beta, \ldots$ есть такова: с помощью равенств:

$$\mathrm{K} = \mathrm{M}(1) + \mathrm{M}(0) = \frac{\mathrm{aK}}{\mathrm{p}} = \frac{bK}{q} = \ldots = \frac{vK}{\alpha} = \frac{wK}{\beta} = \ldots,$$

где, по приведении всех многочленов к дисъюнктному виду, символы $a, b, c, \ldots, v, w, \ldots$ принимаются алгебраическими, исключить все эти символы из пары неравенств:

$$P(u) < \frac{M(1)}{K}, \quad P(u) > \frac{M_0(0)M(1)}{K},$$

в которых тоже все многочлены должны быть приведены к дисъюнктному виду, символы же $a, b, c, \ldots, v, w, \ldots$ трактуются количественными.

Таков общий способ решения задачи, формулированной в начале статьи. Как видим, вообще для искомой вероятности $P(u)$ получаются только пределы, между которыми она содержится; и только тогда, когда

$$M_0(0)M(1) = M(1),$$

получается точное определение $P(u)$, именно:

$$P(u) = \frac{M(1)}{K}.$$

§ 14. Обращаемся к примерам.

Пример 1-й. Пусть вероятность, что умрет в таком-то году или ***A***, или ***B***, (или оба), есть *p*; вероятность, что не умрет в том же году или ***A***, или ***B*** (или оба), есть *q*. Найти вероятность, что умрет в том же году только один из них (т. е. или ***A***, при чем ***B*** останется жив, или обратно).

Пусть *x* событие смерти ***A***, *y* - смерти ***B***.

Даны: $P(x + y) = p$, $P(x_o + y_o) = q$. Ищется $P(xy_o + x_o y)$.

Здесь мы имеем 3 функции. Положим:

$$x + y = s, \quad x_o + y_o = t, \quad xy_o + yx_o = w.$$

Задачу можно считать содержащей пять простых классов, связанных этими тремя условиями, или, что тоже, одним следующим:

$$1=[s(x+y)+sx_oy_o][t(x_o+y_o)+t_oxy]\cdot$$
$$\cdot[w(xy_o+x_oy)+w_o(x_oy_o+xy)]=$$
$$=stwxy_o+stwx_oy+st_ow_oxy+s_otw_ox_oy_o\ .$$

Нам надо найти из этого равенства выражение для w черезъ s и t; лишние классы x и y надо исключить (что достигается подстановкою: $x=1,\ y=1,\ x_o=1,\ y_o=1$). Результат этого исключения есть:

$$1=M(s,t,w)=stw+st_ow_o+s_0tw_o=M(w),$$

откуда

$$M(1)=st,\ \ M(0)=st_o+s_ot,\ M_o(0)=st+s_ot_o\ ,$$
$$M_o(0)M(1)=st,\ \ k=M(1)+M(0)=s+s_ot,\quad ,$$
$$Ks=s,\ \ Kt=ts+ts_o=t.$$

Так как в данном случае $M_o(0)M(1)$ равно $M(1)$, то два неравенства, определяющие функцию w, сводятся на одно равенство

$$w=M(1)=st\ .$$

И действительно, произведение $s=x+y$ на $t=x_o+y_o$ есть $w=xy_o+x_oy$.

И так, искомая вероятность $P(w)$ определится равенством

$$P(w)=\frac{M(1)}{K}=\frac{st}{s+s_0t},$$

после исключения из него, считаемых количественными, символов s и t с помощью равенств:

$$K=s+s_ot=\frac{s}{p}=\frac{t}{q}\ .$$

Из этих равенств имеем:

$$p=\frac{s}{K},\ \ q=\frac{t}{K},\ \ p+q=\frac{s+t}{K},\ \ p+q-1=\frac{s+t-K}{K}=$$
$$=\frac{s+t-(s+(1-s)t)}{K}=\frac{t-t+ts}{K}=\frac{ts}{K}.$$

Следовательно, окончательно:

$$P(w)=p+q-1\ .$$

Для поверки заметим следующее. Если $P(x+y)=p$ то $P[(x+y)_o]=P(x_oy_o)=1-p$. Точно так же, если $P(x_o+y_o)=q$, то $P(xy)=1-q$.

След.

$$P(xy + x_o y_o) = P(xy) + P(x_o y_o) = 2 - p - q\ ,$$

а потому

$$P(xy_o + x_o y) = P[(xy + x_o y_o)_o] = 1 - [2 - p - q] = p + q - 1\ ,$$

результат, вполне согласный с найденным выше.

Пример 2-ой. Пусть вероятность, что свидетель ***A*** показывает истину, есть *p*; вероятность, что свидетель ***B*** показываете истину, есть *q*; вероятность несовпадения их показаний есть *r*. Найти вероятность, что если их показания совпадают, то получается истина.

Пусть классы случаed, когда свидетели *A* и *B* соответственно показывают истину, суть *x* и *y*. Даны:

$$P(x) = p,\ P(y) = q,\ \ P(xy_o + x_o y) = r\ .$$

Ищется отношение

$$\frac{P(xy)}{P(xy + x_o y_o)} = \frac{P(xy)}{1 - r}\ .$$

Очевидно, достаточно найти только $P(xy)$ $P(xy)$ посредством *p*, *q* и *r*. Пусть

$$xy_o + x_o y = s,\quad xy = w\ .$$

Совокупность этих двух условий равнозначна с одним равенством:

$$1 = ws_o xy + w_o(sx_o y + s_o x_o y_o + sxy_o)\ .$$

Вот какому условию подчинена данная задача о четырёх простых классах *x*,*y*,*s*,*w*. Требуется определить *w* через все прочие классы.

Имеем:

$$1 = ws_o xy + w_o(sx_o y + s_o x_o y_o + sxy_o) = M(w),$$

$$M(1) = s_o xy,\ \ M(0) = s(x_o y + xy_o) + s_o x_o y_o,$$

$$M_o(0) = s(xy + x_o y_o) + s_o(x + y),\ \ M_o(0)M(1) = s_0 xy = M(1),$$

$$K = M(1) + M(0) = s_o xy + s_o x_o y_o + sx_o y + sxy_o.$$

Так как $M_o(0)M(1) = M(1) = s_o xy$, то, вместо двух неравенств, *w* определяется одним равенством

$$w = s_o xy\ .$$

Кроме того,

$$Kx = s_o xy + sxy_o,\ \ Ky = s_o xy + sx_o y,\ Ks = sx_o y + sxy_o\ .$$

Считая *x*, *y* и *s* количественными символами, нам надо исключить их из формулы:

$$P(w) = \frac{s_o xy}{K}$$

с помощью отношений

$$\frac{xys_0 + xy_0s}{p} = \frac{xys_0 + x_0ys}{q} = \frac{x_0ys + xy_0s}{r} = K = s_0xy + s_0x_0y_0 + sx_0y + sxy_0.$$

Имеем:

$$r = \frac{sx_oy}{K} + \frac{sxy_o}{K},$$

$$q = \frac{s_oxy}{K} + \frac{sx_oy}{K},$$

$$p = \frac{s_oxy}{K} + \frac{sxy_o}{K} = \frac{s_oxy}{K} + (r - \frac{sx_oy}{K}) = \frac{s_oxy}{K} + r + \frac{s_oxy}{K} - q.$$

След.

$$\frac{s_oxy}{K} = \frac{p+q-r}{2}.$$

А потому окончательно:

$$P(w) = \frac{p+q-r}{2},$$

$$\frac{P(xy)}{P(xy + x_oy_o)} = \frac{p+q-r}{2(1-r)}.$$

Пример 3-й. Пусть из наблюдений относительно эпидемий в какой-нибудь местности найдено, что p есть вероятность посещения отдельного дома горячкой, q - холерой, r есть вероятность непосещения дома обеими болезнями при удовлетворительности санитарных его условий.

Найти вероятность неудовлетворительности санитарных условий отдельного дома в той же местности.

Пусть x - посещение дома горячкой, y холерой, z - неудовлетворительность санитарных условий дома. Даны:

$$P(x) = p,\ P(y) = q,\ P(x_0y_oz_0) = r.$$

Найти $P(z)$. Пусть

$$x_0y_oz_0 = w.$$

Условие, которому подчинена данная задача о четырех простых классах x,y,z,w, есть:

$$1 = wx_oy_oz_o + w_o(x + y + z) = F(z).$$

Отсюда надо найти z посредством x,y,w.

Имеем:

$$F(1) = w_o,\ F(0) = wx_oy_o + w_o(x + y)$$

$$F_o(0) = w(x + y) + w_ox_oy_o,\ F_o(0)F(1) = w_ox_oy_o.$$

Следовательно

$$z < w_o,\ \ z > w_o x_o y_o.$$

Кроме того,

$$K = F(1) + F(0) = w_o + w x_o y_o + w_o (x + y) = w_o + w x_o y_o,$$

$$Kx = x w_o,\ \ ky = y w_o,\ Kw = w x_o y_o.$$

Надо исключить, считаемые количественными, символы *w,x,y* из неравенств

$$P(z) < \frac{w_0}{K},\quad P(z) > \frac{w_o x_o y_o}{K}$$

с помощью отношений:

$$\frac{x w_0}{p} = \frac{y w_0}{q} = \frac{w x_0 y_0}{r} = K = w_o + w x_o y_o.$$

Имеем:

$$w_o = \frac{w x_o y_o}{r} - w x_o y_o = \frac{w x_o y_o (1-r)}{r} = K(1-r);\ \frac{w_o}{K} = 1 - r\ ;$$

$$p + r = \frac{x w_o + w x_o y_o}{K};\ 1 - p - r = \frac{K - x w_o - w x_o y_o}{K} = \frac{w_o - x w_o}{K} = \frac{x_o w_o}{K};$$

$$q + r = \frac{y w_o + w x_o y_o}{K};\ 1 - q - r = \frac{K - y w_o - w x_o y_o}{K} = \frac{w_o - y w_o}{K} = \frac{y_o w_o}{K};$$

$$(1 - p - r)(1 - q - r) = \frac{w_o^2 x_o y_o}{K^2};$$

$$\frac{(1 - p - r)(1 - q - r)}{1 - r} = \frac{w_o^2 x_o y_o}{K^2} \cdot \frac{K}{w_o} = \frac{w_o x_o y_o}{K}.$$

А потому окончательно:

$$P(z) < 1 - r,\quad P(z) > \frac{(1 - p - r)(1 - q - r)}{1 - r}.$$

Пример 4-й. Пусть относительно шаров, находящихся в данной урне, известно, что всякий белый шар есть или крупный, или немраморный. Пусть при вынимании шаров из этой урны обращается внимание только на такие случаи, когда вынутый шар есть или белый, или крупный, или мраморный. Пусть при этих условиях найдено для вероятности случая, когда вынутый шар есть и белый, и крупный, число *p*. Найти вероятность, что будет вынут шар или белый, но некрупный, или, если не белый, то или крупный, или же мраморный.

Пусть *x* - вынутие белого шара, *y* - крупного, *z* - мраморного.

Первоначальные два условия задачи суть:

$$\begin{matrix} x = x(y + z_o) \\ 1 = x + y + z. \end{matrix}.$$

Дана вероятность $P(xy) = p$. Ищется вероятность $P(xy_o + x_o(y+z))$.

Положим

$$xy = u,\ \ xy_o + x_o(y+z) = v.$$

Можно сказать, что данная задача содержит 5 простых классов: *x,y,z, u, v*, подчиненных всем, написанным выше, четырем условиям. Все эти условия совмещаются в одно следующее:

$$1 = [x_o + y + z_o][x + y + z][uxy + u_o x_o + u_o y_o] \times$$
$$\times [vxy_o + vx_o y + vx_o z + v_o xy + v_o x_o y_o z_o] =$$
$$= uv_o xy + u_o vx_o y + u_o vxz_o + u_o vxy_o z_o.$$

По смыслу задачи, отсюда требуется определить *v* через *u*. Лишние классы: *x,y,z* должны быть исключены, что достигается подстановкою:

$$x = y = z = x_o = y_o = z_o = 1.$$

По исключению получим:

$$1 = uv_o + u_o v = F(v).$$

Отсюда имеем:

$$F(1) = u_o,\ F(0) = u,\ \ F_o(0) = u_o,\ \ F_o(0)F(1) = u_o.$$

След. в данном случае *v* определяется равенством

$$v = u_o.$$

Далее, имеем:

$$K = F(1) + F(0) = u_0 + u = 1.$$

Следовательно, условие $1 = K$, которому подчинена функция *u*, сводится на тождество $1 = 1$, что равнозначно с отсутствием всякого условия. А потому получим окончательно:

$$P(v) = P(u_0) = 1 - p.$$

ПРИЛОЖЕНИЕ

О НУМЕРАЦИИ ЛОГИЧЕСКИХ РАВЕНСТВ ВООБЩЕ

Выше (§ 1) было показано, что каждому логическому равенству

$$f(a,b,c,...) = \varphi(a,b,c,...). \qquad (1)$$

соответствует количественное равенство:

$$N[f(a,b,c,...)] = N[\varphi(a,b,c,...)], \qquad (2)$$

выражающее равенство чисел предметов, содержащихся в классах *f* и φ.

Чрез деление обеих частей этого последнего равенства на число $N(1)$, означающее число предметов мира речи, получается еще одно числовое равенство

$$P[f(a,b,c,...)] = P[\varphi(a,b,c,...)], \qquad (3)$$

выражающее равенство вероятностей логических классов *f* и φ.

Назовем для краткости переход от равенства (1) к равенству (3) *пробабилизацией* логического равенства (1); переход же от равенства (1) к равенству (2)—*нумеризацией* логического равенства (1).

Выше мы занимались *непосредственной* пробабилизацией логического равенства, делая переход прямо от равенства (1) к равенству (3), без посредства промежуточного равенства (2). При этом для установления свойств символа ***P*** нам было необходимо пользоваться некоторыми истинами Теории Вероятностей.

Но если мы построим правила для перехода от равенства (1) к равенству (2), причем при установлении свойств символа ***N*** уже нельзя будет пользоваться истинами Теории Вероятностей, то, в виду простой связи между равенствами (2) и (3), в правилах этих мы получим вместе с тем новый способ определения некоторых свойств символа *P*.

Следует также заметить, что равенство (2) может иметь значение не только в качестве промежуточного между (1) и (3), но и само по себе, так как оно может найти себе применение в других областях знаний, напр. в Статистике.

Обращаюсь к построению правил нумеризации логических равенств.

Для нумеризации логич. равенства достаточно нумеризировать каждую его часть порознь и затем приравнять между собою результаты. Таким образом, нумеризация логических равенств сводится к нумеризации отдельных логических функций.

Определение числа предметов, содержащихся в каждом логич. классе *a*, т. е. числа $N(a)$, может быть достигнуто с помощью непосредственного их счета на самом деле. Однако, зная зависимость между некоторыми из символов $N(a)$, $N(b)$, $N(a+b)$, $N(ab)$ и пр., мы можем определять величину одних из этих символов по данным величинам других.

Установление разных видов зависимостей между различными символами ***N*** и составляет предмет теории нумеризации.

Найдем сначала отношение между двумя символами $N[f_o(a,b,c)]$ и $N[f(a,b,c)]$, где f_0 есть логическое отрицание *f*.

Из логического тождества

$$f(a,b,c,...) + f_o(a,b,c,...) = 1$$

имеем:

$$N[f(a,b,c,...)]+ f_o(a,b,c,...)] = N(1).$$

Но так как произведение ff_o равно нулю, то все предметы функции f отличны от предметов функции f_0, а потому

$$N[f + f_o)] = N(f) + N(f_o).$$

и следовательно

$$N(f) + N(f_o) = N(1),$$

откуда

$$N[f_o(a,b,c,...)] = N(1) - N[f(a,b,c,...)].$$

Это и есть искомое отношение. Деля в нём обе части на $N(1)$, получим отношение

$$P[f_o(a,b,c,...)] = 1 - P[f(a,b,c,...)],$$

т. е. одну из основных истин Теории Вероятностей.

Найдем выражение для символа $N(a+b)$. Если a и b дисъюнктны, т. е. $ab = 0$, то понятно, что

$$N(a+b) = N(a) + N(b).$$

Но пусть a и b конъюнктны. т. е. ab отлично от нуля. Из логического тождества

$$a = ab + ab_o,$$

где в правой части оба члена дисъюнктны, получаем

$$N(a) = N(ab) + N(ab_o).$$

Точно так же из тождества

$$b = ab + a_o b,$$

где опять оба члена правой части дисъюнктны, находим

$$N(b) = N(ab) + N(a_o b).$$

Складывая выражения для $N(a)$ и $N(b)$, будем иметь:

$$N(a) + N(b) = 2N(ab) + N(ab_o) + N(a_o b).$$

С другой стороны, сумма предыдущих выражений для a и b доставляет, нам (на основании общего закона логики $m + m = m$) логическое равенство:

$$a + b = ab + ab_o + a_o b,$$

в котором в правой части все три члена дисъюнктны друг с другом. А потому

$$N(a+b) = N(ab) + N(ab_o) + N(a_o b).$$

Сравнение этого выражения с найденным выше показывает, нам, что вообще

$$N(a+b)=N(a)+N(b)-N(ab)\,,$$

откуда, в частности, для случая, когда $ab=0$ и след. $N(ab)=N(0)=0$, получим, как и ранее,

$$N(a+b)=N(a)+N(b)\,.$$

Далее, легко видеть, что вообще (в силу доказанного, а также закона $mm=m$):

$$N(a+b+c)=N[(a+b)+c]=N(a+b)+N(c)-N[(a+b)c]=$$
$$=N(a)+N(b)-N(ab)+N(c)-N[ac+bc]=$$
$$=N(a)+N(b)+N(c)-N(ab)-[N(ac)+N(bc)-N(abc)]=$$
$$=[N(a)+N(b)+N(c)]-[N(ab)+N(ac)+N(bc)]+N(abc).$$

Точно так же мы нашли бы:

$$N(a+b+c+\partial)==N(a)+N(b)+N(c)+N(\partial)]-$$
$$-[N(ab)+N(ac)+N(a\partial)+N(bc)+N(b\partial)+N(c\partial)]+$$
$$+[N(abc)+N(ab\partial)+N(bc\partial)]-N(abc\partial).$$

Закон построения этих формул очевиден. В частности, когда все слагаемые классы дисъюнктны между собою, мы находим:

$$N\Sigma a^{(i)}=\Sigma N(a^{(i)})\,.$$

откуда, по разделении на $N(1)$, получаем отношение:

$$P(a'+a''+a'''+...)=P(a')+P(a'')+P(a''')+...\,,$$

т.е. еще одну истину Теории Вероятностей, на которую мы ссылались выше.

Можно найти иное выражение для символа $N\Sigma a^{(i)}$.

Так как в логике имеет место тождество:

$$a'+a''=a'+a'_{o}\,a''\,,$$

где в правой части оба члена дисъюнктны между собою, то

$$N(a'+a'')=N(a')+N(a'_{o}\,a'')\,.$$

Далее, зная, что

$$a'+a''+a'''=a'+a'_{o}\,a''+a'_{o}\,a''_{o}\,a'''\,,$$

где опять все члены правой части дисъюнктны между собою, найдем:

$$N(a'+a''+a''')=N(a')+N(a'_{o}\,a'')+N(a'_{o}\,a''_{o}\,a''')\,.$$

Точно так же найдем и вообще:

$$N(a'+a''+a'''+...)=N(a')+N(a'_{o}\,a'')+N(a'_{o}\,a''_{o}\,a''')+...\,.$$

Третий прием для определения символа $N\Sigma a^{(i)}$ заключается в разложении суммы $\Sigma a^{(i)}$ на элементы объема (которые всегда дисъюнктны между собою). Поэтому, если такое разложение есть:

$$\Sigma a^{(i)} = s'+s''+s'''+...,$$

то понятно, что

$$N\Sigma a^{(i)} = N\Sigma s^{(k)}.$$

Наконец, четвертый прием определения того же символа есть следующий. Так как отрицание суммы $a'+a''+a'''+...$ есть произведете $a'_o\, a''_o\, a'''_o...$ то понятно, что

$$N(a'+a''+a'''+...) = N(1) - N(a'_o\, a''_o\, a'''_o...)\,.$$

Обращаемся к определению символа $\boldsymbol{N}$ от произведения логических классов. Выше было доказано, что

$$N(a+b) = N(a) + N(b) - N(ab)\,,$$

а потому

$$N(ab) = N(a) + N(b) - N(a+b)\,.$$

Легко также видеть, что

$$N(ab) = N(1) - N[(ab)_o] = N(1) - N(a_o + b_o)...\,. \qquad (\text{E})$$

Обобщением этих формул я заниматься не буду. Вместо того, обращу внимание на следующее. Предпоследняя формула показывает нам, что, зная символы $N(a)$ и $N(b)$, мы еще не можем определить величины символа $N(ab)$. Однако, легко указать пределы, внутри которых содержится величина этого символа; а именно: $N(ab)$ не меньше нуля и не больше наименьшего из символов $N(a)$ и $N(b)$.

Буль доказал, что нижний из этих пределов можно формулировать обстоятельнее. А именно, он доказывает, что $N(ab)$ не меньше

$$N(a) + N(b) - N(1)\,.$$

В самом деле, из формулы (E) следует, что

$$N(ab) = N(1) - N(a_o + b_o) = N(1) - [N(a_o) + N(b_o) - N(a_o b_o)] =$$
$$= N(1) - [N(1) - N(a) + N(1) - N(b) - N(a_o b_o)] = \qquad .$$
$$= N(a) + N(b) - N(1) + N(a_o b_o).$$

Вот новое выражение для символа $N(ab)$, откуда видим, что, действительно, $N(ab)$ не меньше

$$N(a) + N(b) - N(1)\,.$$

Далее, для случая трех множителей имеем:

$$N(a'a''a''') = N[(a'a'')a'''] = N(a'a'') + N(a''') - N(1) +$$
$$+ N((a'_o + a''_o)a'''_o) =$$
$$= N(a') + N(a'') - N(1) + N(a'_o a''_o) + N(a''') - N(1) + \quad .$$
$$+ N((a'_o + a''_o)a'''_o) = N(a') + N(a'') + N(a''') - 2N(1) +$$
$$+ [N(a'_o a''_o) + N((a'_o + a''_o)a'''_o)].$$

Так как каждый из символов $\boldsymbol{N}$ не меньше нуля, то отсюда следует, что $N(a'a''a''')$ не меньше

$$N(a') + N(a'') + N(a''') - 2N(1).$$

Такими же суждениями можно доказать, что вообще

$N(a'a''a'''...a^{(m)})$ не меньше $\Sigma N(a) - (m-1)N(1)$.

Таков нижний передел величины символа $\boldsymbol{N}$ от произведения классов. Что же касается верхнего, то понятно, что величина того же символа не больше величины наименьшего из символов $N(a'), N(a''),..., N(a^{(m)})$..

Вот собственно и все, что мне известно о правилах нумеризации.

В заключение замечу следующее. Выше мы получили из правил нумеризации две основные истины Теории Вероятностей. Однако, дальнейшие истины той же науки мы можем получить из правил нумеризации только при помощи гипотезы о равномерном распределении предметов каждого класса по всему протяжению мира речи. Например, только при условии этой гипотезы, мы можем сказать, что $N(ab)$ составляет, ту же часть от $N(a)$, как $N(b)$ от $N(1)$, т. е. написать пропорцию.

$$N(ab) : N(a) = N(b) : N(1),$$

откуда

$$N(ab) = \frac{N(a)N(b)}{N(1)},$$

и следов., по разделении на $N(1)$:

$$P(ab) = \frac{N(a)}{N(1)} \cdot \frac{N(b)}{N(1)} = P(a)P(b).$$

УДК 519.6 + 53.088

С.Ф. Свиньин, А.И. Попов

ФИНИТНЫЕ БАЗИСНЫЕ ФУНКЦИИ В ЗАДАЧАХ ФОРМИРОВАНИЯ ВЫБОРОК СИГНАЛОВ КОНЕЧНОЙ ПРОТЯЖЕННОСТИ

Свиньин С.Ф., Попов А.И. **Финитные базисные функции в задачах формирования выборок сигналов конечной протяженности.**

Аннотация. В статье рассматриваются вопросы применения систем базисных функций, определенных на конечных интервалах аргументов, в задаче формирования дискретных выборок сигналов. Такие базисы позволяют обосновать объемы сеток выборок реальных сигналов при ситуациях, когда их спектры инфинитны и характеризуются определенной степенью затухания в области высоких частот. Для финитных функциональных зависимостей, у которых аргументом не является время, теряет смысл понятие частоты Найквиста.

Ключевые слова: инфинитный спектр, дискретная выборка, конечная энергия сигнала, компактный носитель, вейвлеты Добеши, быстрые спектральные преобразования.

Svinyin S.F, .Popov A.I. **Finite Basic Functions in the Tasks of Sampling Signals of Finite Qxtension.**

Abstract. The article deals with the application of systems of the basic functions, defined on finite argument intervals, in the problem of obtaining discrete signal samples. These mathematical bases allow justifying the size of signal sample lattices for actual situations where their spectra are infinite and are characterized by a certain degree of attenuation at high frequencies. For expressions with finite functions, which do not have the time as argument, the conception “the Nyquist frequency” loses its significance.

Keywords: signal, infinite spectrum, sampling, finite energy, compact carrier, Daubechies wavelets, fast spectral transforms.

Введение. В 2015 году исполнилось 100 лет со дня опубликования Э. Уиттекером статьи, в которой детально исследовался так называемый «кардинальный ряд» в математике [1]. Автор дал общему члену $sin(\pi x)/(\pi x)$ этого ряда обозначение $sinc(x)$, и существенным было то, что аргумент функции рассматривался как абстрактная математическая переменная, не связанная с какими-либо физическими величинами. Во главу угла был положен принцип финитности полосы спектра аналитических целых функций, и для абсолютно точного их восстановления требовалась фильтрация с помощью идеальных фильтров низкой частоты (ФНЧ). Предполагалось, что эти фильтры обладают частотной характеристикой в виде rect-функции, и абсцисса точки ее разрыва впоследствии получила впоследствии название частоты среза и обозначалась как ω_c.

С 30-40-х годов XX века кардинальный ряд получил широкое распространение в теории связи после того, как В.А. Котельников и К. Шеннон дали формулировки теорем отсчетов [2,3]. С 50-х годов в мировой литературе появилось большое число работ, опиравшихся при

анализе сигналов – функций времени на принцип финитности спектров. Возникли проблемы фактора усеченности ряда, неидеальной фильтрации, интерполяционного критерия восстановления функции и т.д. Важный шаг был сделан, когда стали применяться энергетические критерии оценок ошибок восстановления. Обзор первой группы отечественных работ по теории выборок дал академик А.А. Харкевич в 1958 году [4]. Он обратил внимание на условность термина «граничная частота полосы» для реальных сигналов и на необходимость расширения теорем отсчетов на область случайных сигналов, которые могут иметь очень широкий и, к тому же, гладкий спектр в области высоких частот. Им же замечено, что для сигналов, длительность которых превышает так называемый интервал корреляции, теорема Котельникова в применении к процессам с неограниченным спектром должна рассматриваться как приближенное утверждение.

В дальнейшем появились возможности строить модели, лишь приближенно отражающие точность первичных данных измерений сигналов. Наиболее распространенные причины ошибок, возникающих в процессе дискретных выборок, были рассмотрены в работе [5]. Очевидно, что усечение ряда практически не ставит под сомнение формулировки теорем отсчетов для функций времени, если длительность сигнала значительно превышает (в десятки, сотни тысяч и более раз) величину периода самой низкочастотной его составляющей. Для функций, аргумент которых не является временем, это условие выполняется далеко не всегда. Функции конечной (финитной) протяженности имеют инфинитные спектры. Этот факт приобретает особое значение для функций нескольких переменных, когда они определены на компактных носителях (ограниченных по размерам площадях, параллелепипедах и т.д.). Создается основа для восстановления по дискретным выборкам финитных сигналов посредством систем финитных базисных функций, свойства которых существенно отличаются от свойств бесконечного кардинального ряда. Значительное распространение в теории аппроксимации, начиная с 70-80-х годов, получили локальные базисные функций, в том числе рассматриваемые на компактных носителях. В этой статье отразим их роль в теории выборок.

2. Финитность сигналов – инфинитность спектров. С 1960-х годов XX века приоритет в исследованиях в рамках теории выборок получил энергетический подход. Обращено внимание на класс функций с конечной энергией, т.е. удовлетворяющих условию [6]:

$$\int_{-\infty}^{+\infty} \left| f(t) \right|^2 dt < \infty. \qquad (1)$$

В краткой заметке [7] задача оценки ошибки восстановления была поставлена следующим образом: имеется «неизвестная» (по их мнению) функция времени $f(t)$, но с ограниченным спектром. Требовалось приближенно восстановить ее на некотором конечном интервале времени $(-T,T)$, зная выборки с шагом $h < (1/\omega_c)$. Авторы вывели аналитическую формулу для модуля разности функции и усеченного кардинального ряда в виде неравенства:

$$\varepsilon = \left| f(t) - \sum_{i=-k}^{k} f(ih) \frac{\sin \frac{\pi}{h}(t - ih)}{\frac{\pi}{h}(t - ih)} \right| \le \frac{\sqrt{2}}{\pi} E \left| \sin\left(\frac{\pi t}{h}\right) \right| \sqrt{\frac{Th}{T^2 - t^2}}, \qquad (2)$$

где энергию E они считали конечной, причем полной и равной спектральной энергии E_c:

$$E = E_c = \int_{-\omega c}^{\omega c} F^2(\omega) d\omega.$$

Различные оценки ошибок, обусловленных отбрасыванием высокочастотной части спектра, рассматривались на протяжении десятков лет во многих других работах. В частности, в статье [8] приводились доказательства наличия значительной вычислительной неустойчивости процесса восстановления при малейшем отклонении произведения $2\omega_c T$ от теоретического значения, как в сторону его уменьшения, так и в сторону увеличения. В монографии отечественных авторов Я.И. Хургина и В.П. Яковлева [6].

В другой широко известной статье [9] предлагался ответ на вопрос о подходе к ограниченности ширины полосы. Подчеркивался тот факт, что «реальные сигналы где-то начинаются и где-то кончаются, и, следовательно, их полоса не может быть ограниченной». В этой работе были введены определения, уточняющие критерии восстановления: «сигнал, ограниченный во времени на уровне ε» и «полоса частот $(-\omega_c,\omega_c)$, ограниченная на уровне ε».

По аналогии с ними введем понятие энергии E_ε, ограниченной на уровне ε как значение интеграла энергии для функций, интегрируемых с квадратом, отличающееся на величину ε от полной энергии E. Также будем иметь в виду равенство Парсеваля для таких функций при $\omega > 0$:

$$E = \int_{-T}^{T} f^2(t) dt = \frac{1}{\pi} \int_{0}^{\infty} F^2(\omega) d\omega, \qquad (3)$$

Наиболее полные анализы многочисленных работ по теории

отсчетов и ее приложений в канун 80-х годов, включая разложения по базисам, отличающимся от кардинального ряда, проделаны в виде больших обзоров, опубликованные в 1977 году [10,11]. Авторы обоих работ проанализировали в сумме около пятисот публикаций. В том числе выделены достижения в области развития подходов к дискретизации функций нескольких переменных (ФНП). В связи с этим вопросы формирования выборок многомерных сигналов, начиная с функций двух переменных, когда фактор инфинитности спектров явно проявляется, кратко затронем в следующем разделе.

3. О выборках сигналов – функций двух и более переменных. Традиционная теория отсчетов одномерных функций с финитным спектром была расширена и на процессы обработки двумерных изображений. Обобщения результатов можно найти в монографиях [13,14]. В первой из них делается предположение, что спектр изображения $F(\omega_x,\omega_y)$ финитен и равен нулю вне определенного прямоугольника $|\omega_x|<=\omega_{cx}$, $|\omega_y|<=\omega_{cy}$. Затем берутся ортогонально расположенные отсчеты с шагами $1/2\omega_{cx}$ и $1/2\omega_{cy}$ и непрерывное изображение $f(x,y)$ выражается через дискретные значения $f(n_1/2\omega_{cx}, n_2/2\omega_{cy})$. Интерполяционная функция двумерной дискретизации получает вид:

$$w(x,y)=\left(\frac{\sin(2\pi\omega_x x)}{2\pi\omega_x x}\right)*\left(\frac{\sin(2\pi\omega_y y)}{2\pi\omega_y y}\right), \qquad (4)$$

где w – интерполяционная функция двумерной дискретизации.

В монографии [13] теорема двумерной дискретной выборки изображений приведена в следующей формулировке: «Если $F(\omega_x,\omega_y)$ имеет носитель, содержащийся в интервале $[-\pi/T_x,\pi T_x]*[-\pi T_y,\pi T_y]$, то:

$$f(x,y)=\sum_{i=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} f(iT_x,kT_y)h_x(x-iT_x)h_y(y-kT_y), \qquad (5)$$

где:

$$h_x=\frac{\sin(\pi x T_x)}{\pi x T_x}, \quad h_y=\frac{\sin(\pi y T_y)}{\pi y T_y}$$ ».

Существуют теоремы многомерных выборок, разработанные для случаев, когда число независимых переменных может быть три и более. Примерами могут служить результаты, полученные Ф. Реза [11], а также Д. Петерсеном и Д. Миддлтоном [15]. В их работах формулировки теорем доказывается с позиций теории функций с финитным спектром, представляющим собой вариант искусственного расширения одномерного спектра на многомерные пространства

волновых чисел. Очевидно, что в приложениях для N-мерных эвклидовых пространств сигналов проблемы выбора нескольких полос фиксированной ширины каждая (т.е. границ множеств волновых чисел $\omega_1,\omega_2,\ldots,\omega_N$) многократно возрастают по сравнению с проблемами для функций одной переменной. Области ненулевых частот фиксируются вокруг начала координат как прямоугольники, параллелепипеды, гиперкубы и т.п. Возникает основное противоречие математической теории непрерывных сигналов, поскольку в реальности финитным сигналам должны соответствовать инфинитные спектры и наоборот. Проблема замены многомерных функциональных зависимостей их цифровыми отображениями авторы пытались решить, в частности, за счет периодического продолжения спектров за пределы геометрических форм основных носителей информации. Но тогда нужно периодически продолжать и сами многомерные функции $f(x_1,x_2,\ldots,x_N)$, что далеко не всегда соответствует картинам реальных физических полей.

Рассмотрим вначале вопрос применения финитных базисных функций на примерах анализа профилей поля магнитной индукции на поверхности Земли, затем оценим всю картину поля на участке площадью в несколько сотен квадратных километров, полученную как результат измерений методом аэромагниторазведки. Графическое изображение данного участка поля приведено на рисунке 1.

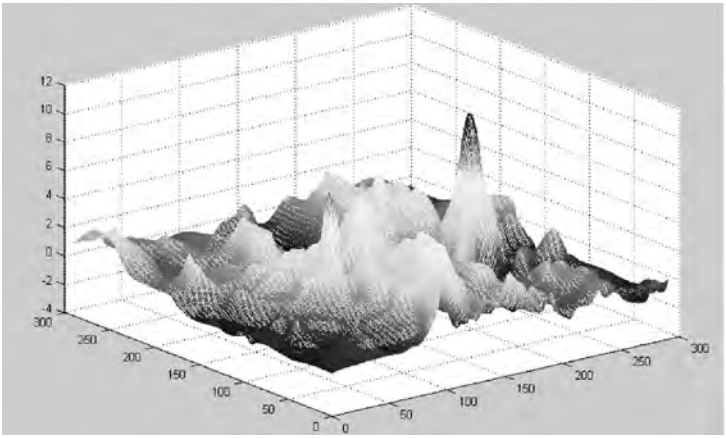


Рис. 1. График поля магнитной индукции на одном из участков поверхности Земли

По осям x,y указаны номера отсчетов, по оси ординат – значения индукции в микроТесла.

4. Базисные функции, заданные на компактных носителях. Среди базисов, заданных на компактных носителях, особое место занимают полиномиальные базисные сплайны (B-сплайнов). С точки зрения применения в теории выборок они интересны, прежде всего, тем, что соответствующие преобразования Фурье элементов аппроксимирующих B-сумм имеют своим результатом простые аналитические выражения [16], напоминающие во многом общий член кардинального ряда Уиттекера-Котельникова-Шеннона. Основное отличие состоит в том, что непрерывным независимым аргументом является частота, а не время:

$$F_B(\omega) = B(0)\left(\frac{\sin\left(\frac{\omega h}{2}\right)}{\frac{\omega h}{2}}\right)^{m+1}. \qquad (6)$$

Здесь h – расстояние между равноотстоящими узлами сплайна, m – степень сплайна. Узлы могут являться опорными точками для выбора оптимальной частоты отсчетов с точки зрения алгоритма минимизации ошибок. График спектральной плотности $F_{B3}(\omega)$ кубического B-сплайна показан на рисунке 2.

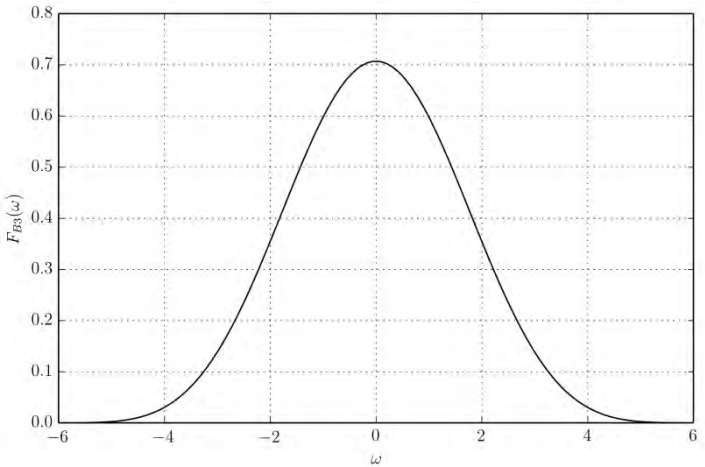


Рис. 2. Спектр одномерного кубического B-сплайна

Рассмотрим функцию $f(x)$ одного аргумента, заданную на замкнутом отрезке $[a,b]$. Известно, что функция достаточной степени гладкости может быть приближенно представлена в виде суммы

«взвешенных» B-сплайнов целой степени m дефекта 1:

$$f(x) \cong \sum_{i=-m}^{n+m} b_i B_i(x), \tag{7}$$

где b_i – коэффициенты. Графики последовательностей B-сплайнов 1-й и 3-й степени показаны на рисунке 3 для случая расстояния между узлами $h=3$.

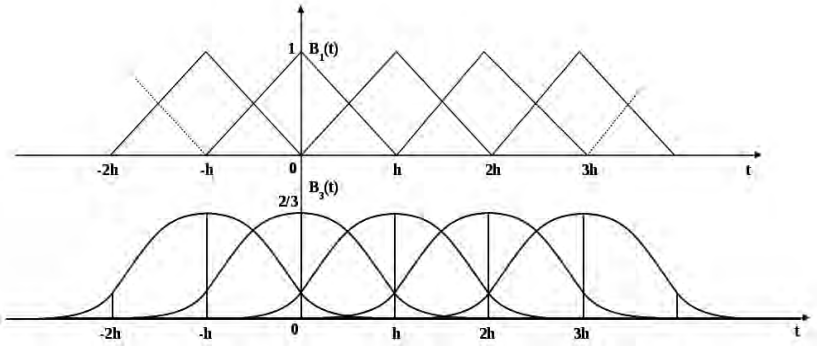


Рис. 3. Последовательности B-сплайнов 1-й и 3-й степени

Введем, кроме использованного выше термина «спектральная энергии сигнала E_c» [6], понятие энергии спектра аппроксимирующей последовательности B-сплайнов $E_{as}(\omega)$. Тогда можно с некоторой степенью приближения записать интегральное соотношение:

$$\int_0^{\infty} |F(\omega)|^2 d\omega \cong \int_0^{\infty} |F_{as}(\omega)|^2 d\omega. \tag{8}$$

Спектры $F(\omega)$ и $F_{as}(\omega)$ инфинитны, но очевидно, что энергия последовательности, заданной на конечном интервале $[a,b]$, при ограничениях на диапазон значений сигнала $f(x)$, конечна. Спектральную энергию как интеграл от квадрата модуля $F_{as}(\omega)$, можно разбить на две части – низкочастотную (НЧ) и высокочастотную (ВЧ):

$$\left(\frac{1}{\pi}\right)\int_0^{\infty} |F_{as}(\omega)|^2 d\omega = \left(\frac{1}{\pi}\right)\int_0^{\omega_e} |F_{as}(\omega)|^2 d\omega + \left(\frac{1}{\pi}\right)\int_{\omega_e}^{\infty} |F_{as}(\omega)|^2 d\omega. \tag{9}$$

Частоту ω_e назовем граничной частотой эффективной ширины полосы НЧ-спектра последовательности. Эта полоса может быть рассчитана по энергии НЧ-части, рассчитанной «с точностью до ε» по

отношению к полной энергии.

На основании теоремы математического анализа об интегральных неравенствах, определим ВЧ-часть энергии последовательности:

$$\varepsilon = \frac{1}{\pi}\int_{\omega_e}^{\infty} \left(F_{\text{as}}(\omega)\right)^2 d\omega < C_1 h^2 \int_{\omega_e}^{\infty} \left(\frac{\sin(\omega h/2)}{\omega h/2}\right)^{2m+2} d\omega < C_1 h^2 \int_{\omega_e}^{\infty} \left(\frac{2}{\omega h}\right) d\omega = \frac{2^{m+2} C_1}{(2m+1)\pi^{2m+1}} h, \qquad (10)$$

где C_1 – коэффициент, зависящий от количества узлов сплайна.

Из данного выражения следует, что энергия высокочастотных составляющих последовательности B-сплайнов, аппроксимирующей непрерывный сигнал $f(x)$, пропорциональна значению шага выборки h с множителем, зависящим от степени сплайна m.

В формуле (10) интервал между узлами, равный шагу выборки $h=2\pi/\omega_e$, может быть рассчитан по значению энергии E ВЧ-части спектра последовательности.

Покажем, что неравенство (10) выполняется, на примере фрагмента одного из профилей $f(x)$ (сплошная линия) магнитного поля, графики которых приведены на рисунке 4. Расстояния между точками измерений по оси абсцисс на практике составляют $h_x=x_{i+1}-x_i=0{,}25$ км.

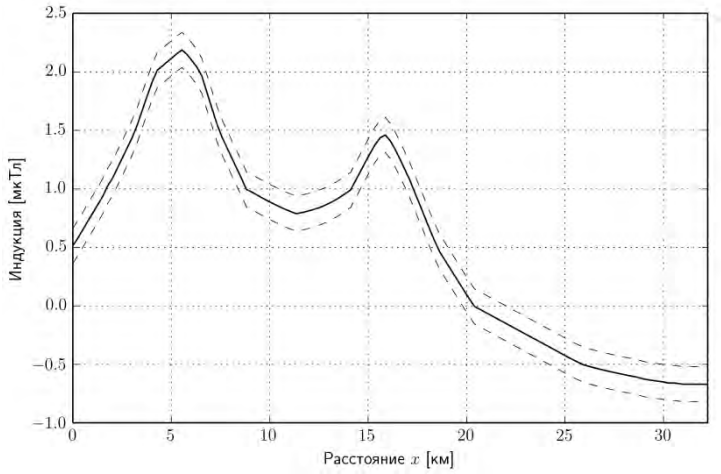


Рис. 4 . Графики профилей поля электромагнитной индукции, наведенной на поверхности Земли

Аппроксимацию сигнала выразим в виде последовательности сглаживающих кубических B-сплайнов в соответствии с формулой

аппроксимации (7). Формулы сглаживания для сплайнов различных степеней разработаны в [17,18]. Аналитическое выражение для модуля амплитудного спектра последовательности B-сплайнов целой степени принимает вид [19]:

$$\left|F_{as}\left(\omega_x\right)\right| = B_0\left|\frac{\sin\left(\omega_x h/2\right)}{\omega_x h/2}\right|^{m+1} * \left|\sum_{i=0}^{n} b_i e^{-j\omega_x h}\right|, \qquad (11)$$

где B_0 – амплитуда B-сплайна с нулевым индексом. Подчеркнем, что аргументом при построении поля является физическое расстояние x вдоль поверхности, и пространственную частоту при выполнении преобразования Фурье обозначим ω_x.

Узлы сплайна x_i для упрощения расчетов, проиндексируем в целых числах: $i = -m,..,0,1,\ldots,n+m$. График спектра при m=3 показан на рисунке 5.

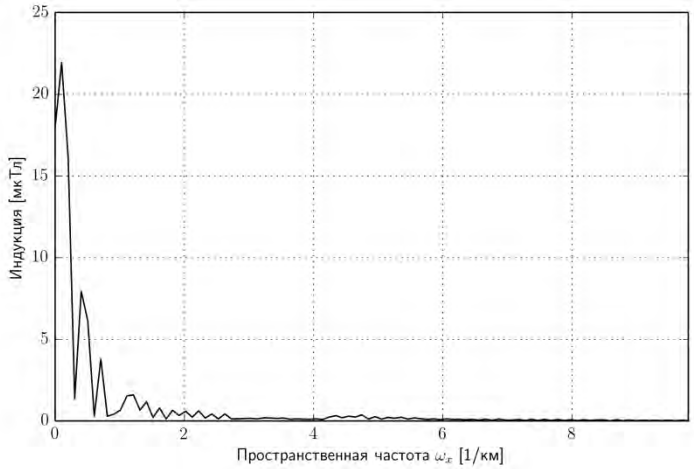


Рис. 5. График спектра аппроксимирующей последовательности кубических B-сплайнов

Значение энергии ВЧ-части спектра как отклонение от полной энергии должно удовлетворять неравенству:

$$\varepsilon(\omega_x) \le \left(\frac{1}{\pi}\right)\int_{\omega_{ex}}^{\omega_x}\left|F_{as}(\omega_x)\right|^2 d\omega_x. \qquad (12)$$

Величина полной энергии спектра $E_c(\omega_x)$ аппроксимирующей

суммы (4) в результате применения формулы:

$$E_c = \left(\frac{1}{\pi}\right)\int_0^{\infty} \left|F_{as}\left(\omega_x\right)\right|^2 d\omega_x, \qquad (13)$$

получилась равной E_c=137,0607.

Если задать отклонение ε от E_c равным, например, на уровне в 0,5%, то величина E_ε = 136,3754. Из равенства (9) получается, что соответствующая относительная граничная частота ω_e эффективной полосы, равна 0,354. Шаг выборки при этом h=2,825, что в пересчете на километры дает число h_x=0,706 км.

При меньшей величине отклонения ε, равной, например, 0,1%, получаем следующие результаты: E_ε=136,924, ω_e=1,280, h=0,781 и в пересчете на километры получаем величину h_x=0,195 км. Это означает, что выбранный геофизиками при проведении реальных измерений шаг h_x выборки точек отсчетов вдоль оси x равен 0,25 км, что соответствует несколько большему значение погрешности, чем 0,1%.

5. Ортонормированные вейвлеты с компактными носителями. Роль алгоритмов быстрых вейвлет-преобразований. Значительный прогресс в использовании вейвлетов в различных приложениях связан, в первую очередь, с наличием быстрых алгоритмов спектральных дискретных преобразований, класс которых значительно шире множества быстрых преобразований в базисе комплексных экспоненциальных функций. Для решения проблемы организации минимальных выборок сигнала, обеспечивающих необходимую точность восстановления, следует провести исследование собственных спектров вейвлет-коэффициентов. Такие системы вейвлет-функций, как производные от функции Гаусса, вейвлеты Морле, вейвлеты Шеннона и др. теоретически определены на всей оси, но могут рассматриваться как локальные. Но основную роль в алгоритмах дискретных быстрых вейвлет-преобразований (БВП) играют ортонормированные вейвлет-базисы, заданные на компактных носителях.

Для применения энергетического критерия точности восстановления сигнала по вейвлет-коэффициентам необходимы два основных оператора: кратномасштабный анализ [14] и вычисление октавного спектра энергии [21]. Достоинством октавного спектра является то, что он, как и спектр Фурье, инвариантен по отношению к сдвигам во времени стационарных сигналов. Свойством кратномасштабного анализа обладают и некоторые вейвлеты, рассматриваемые на всей оси $t \in (-\infty,\infty)$, например, вейвлеты

Шеннона [22, 23].

Преобразуем непрерывный сигнал $f(x)$ к дискретному виду – представим его как вектор-строку, содержащий n действительных чисел f_i, i=0,1,…, n-1. В алгоритмах быстрых вейвлет-преобразований фактически используются целочисленные итерации одного единственного оператора масштабирования D_σ (σ>1), описывающего растяжение [20]. Обычно используется масштаб σ=2, при котором материнский вейвлет удовлетворяет тождеству:

$$D_2\psi(t) = \sum_{k=0}^{n-1} c_k \psi(t-k). \qquad (14)$$

Для функций $f \in L^2(R)$ частная сумма с вейвлет-коэффициентами c_k интерпретируется как разность между двумя приближениями $f - c$ разрешениями 2^{-j+1} и 2^{-j}, и кратномасштабный анализ использует наборы сеток приближения. Приближение с разрешением 2^{-j} содержит всю необходимую информацию для вычисления с более грубым разрешением 2^{-j-1}.

На рисунке 6 приведен граф быстрого преобразования Хаара (БПХ) на n=2^3=8 отсчетов с добавлением операторов вычисления составляющих октавного спектра.

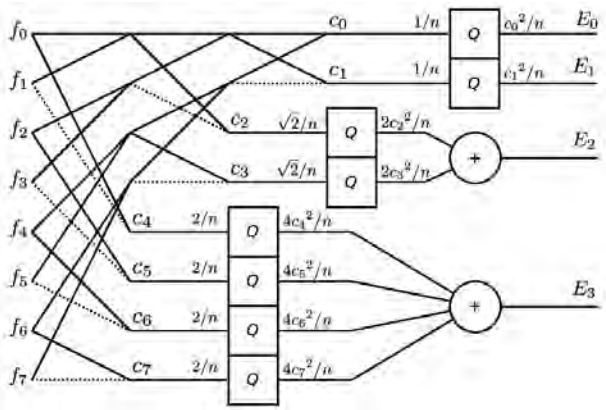


Рис. 6. Граф быстрого преобразования Хаара: c_k – коэффициенты Хаара, Q – квадраторы, E_s – значения октавных составляющих спектральной энергии дискретного сигнала

Авторами было разработано программное обеспечение на языке программирования Python3 с использованием библиотеки Scipy для вычисления значений вейвлет-коэффициентов. На рисунке 7 приведена

гистограмма части значений коэффициентов быстрого преобразования Хаара (БПХ) исходного вектора $\{f_i\}$, содержащего $n=2^p=64$ отсчета отсчетов. Говорят, что «сигнал имеет длину 64». Показатель степени p, означающий максимальное число итераций, носит название порядка дискретного преобразования.

Величина интегральной по всем октавам спектральной энергии вектора коэффициентов Хаара $\{c_k\}$, вычисляется квадратичная сумма о вида [21]:

$$E_\varepsilon = ((c_0^2 + c_1^2) + 2^{-1}\left(c_2^2 + c_3^2\right) + 2^{-2}\sum_{k=4}^{2^{p-4}-1} c_k^2 + 2^{-3}\sum_{k=8}^{2^{p-3}-1} c_k^2 + \ldots + 2^{-p}\sum_{k=2^{p-1}}^{2^p-1} c_k^2)n. \qquad (15)$$

Ее значение получается равным E_ε=135,297.

Выполним быстрое преобразование Хаара повторно для сетки отсчетов, в 2 раза более частой, т.е. для длины сигнала в 128 отсчетов для того же самого отрезка. В этом случае значение энергии равно E_ε= 137,248.

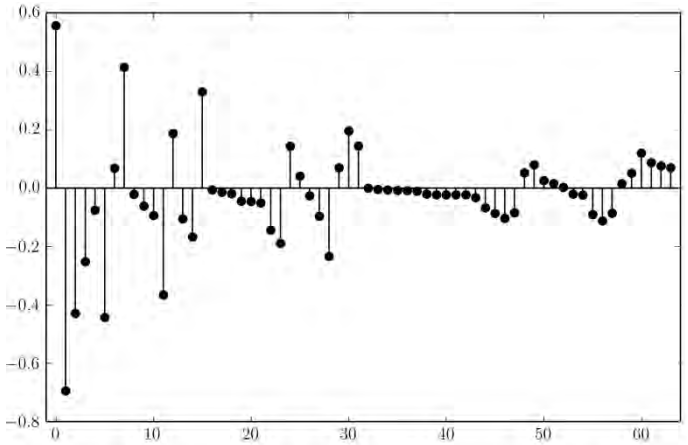


Рис. 7. Значения вейвлет-коэффициентов Хаара функции, график которой изображен на рисунке 4

На рисунке 8 показан график, а на рисунке 9 гистограмма вейвлета Добеши Db4. Значение спектральной энергии, определяемое в результате быстрого вейвлет-преобразования, как и в случае БПХ,

вычисляется как квадратичная сумма по всем октавам вектора коэффициентов Добеши $\{c_k\}$:

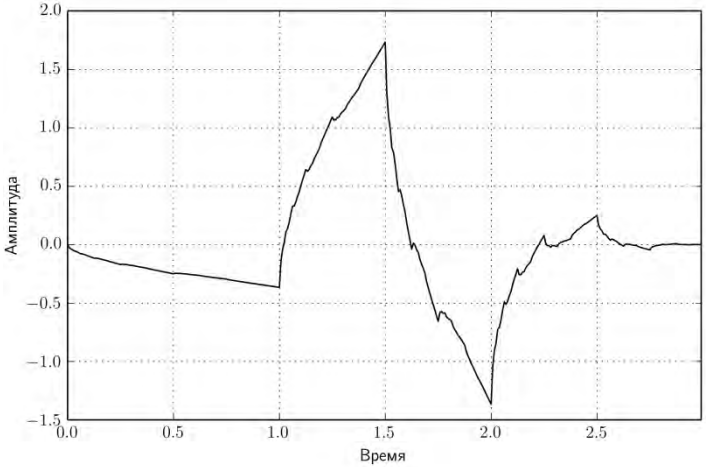


Рис. 8. График материнского вейвлета Db4.

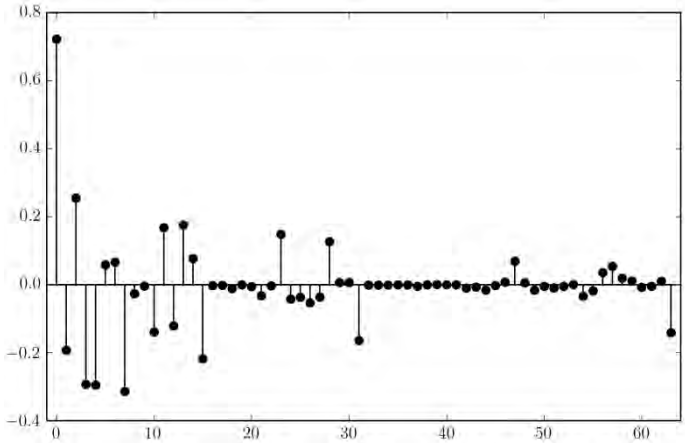


Рис. 9. Гистограмма первых, наиболее значащих по модулю коэффициентов разложения функции $f(x)$ по базисным функциям семейства Db4

$$E_\varepsilon = \left(\sum_{k=0}^{3} c_k^2 + 2^{-1} \sum_{k=4}^{2^{p-4}-1} c_k^2 + 2^{-2} \sum_{k=2^{p-4}}^{2^{p-3}-1} c_k^2 + \ldots + 2^{p} \sum_{k=2^{p-1}}^{2^{p}-1} c_k^2 \right) n. \qquad (16)$$

Когда разрешение 2^{-j} увеличивается, то аппроксимация

совокупности коэффициентов сходится к исходному сигналу $f(x)$ [13]. Аналогично, сходимость последовательности энергий E_ε при увеличении номера итерации j может быть установлена по неравенству

$$\left|E_{\varepsilon,j+1} - E_{\varepsilon,j}\right| \le \varepsilon,\ j = 1,2,... \qquad (17)$$

В частности, при длине дискретного сигнала 2^p=64 E_ε=137, 158, а при длине, равной 128, она составляет E_ε= 137,498 .

6. Заключение. В статье развивается подход к проблеме выборок непрерывных сигналов с позиций теории функций с конечной энергией, причем обладающих инфинитным спектром. Существуют две модели, аналитически описывающие процессы убывания спектра в высокочастотной области: одна - соответствующая гиперболическому закону $F(\omega)\sim\omega^{-m}$, другая – экспоненциальному закону с основанием a: $F(\omega)\sim a^{-\omega}$ [4]. В качестве примеров, подтверждающих возможности энергетического подхода к расчету необходимого шага выборки финитных функций с интегрируемым квадратом, приведены варианты моделей, использующих базисные функции с компактными носителями - B-сплайны и вейвлеты. Сходимость алгоритмов быстрых вейвлет-преобразований к сигналу $f(x)$ в пространстве C_∞ при увеличении числа итераций j=1,2,3, … доказана в [14].

Роль таких моделей и финитных базисов должна возрастать при переходе к алгоритмам получения дискретных выборок многомерных непрерывных сигналов – функций вида $f(x_1,x_2,\ldots,x_N)$, примером которых является поле, показанное на рис.1. Важным свойством локальных ортогональных базисов в пространстве многомерных сигналов, а также их спектров является свойство сепарабельности. Об энергетическом подходе к проблеме выборок применительно к финитным сигналам – функциям двух и более переменных авторы намерены написать следующую статью.

Литература

1. *Whittaker E.* On the Functions which are represented by Expansions of the Interpolation Theory // Proc. Roy. Soc. Edinburgh. 1915. vol. 35. pp. 181–194.
2. *Котельников В.* О пропускной способности «эфира» и проволоки в электросвязи // Успехи физических наук (Приложение). 2006. Т. 176. №7. С. 762–770.
3. *Shannon C.* Communications in the Presence of Noise // Proc. IRE. 1949. Vol. 37. no. 1. pp. 10–21.
4. *Харкевич А.А.* О теореме Котельникова // Радиотехника. 1958. Т.1. №8. С. 3–10.
5. *Папоулис А.* Анализ ошибок в теории выборок // ТИИЭР. 1966. Т.54. №7. С. 34–43.
6. *Хургин Я.И., Яковлев В.П.* Финитные функции в физике и технике // М.: Наука. 1971. 408 с.
7. *Цыбаков Б.С., Яковлев В.П.* О точности восстановления функции с помощью конечного числа членов ряда Котельникова // Радиотехника и электроника. 1959. Т.4. № 3. С. 542.
8. *Ландау Г.* Метод выборок, передача информации и частота Найквиста // ТИИЭР. 1967. Т.55. №10. С.56–62.

9. *Слепян А.Д.* О ширине полосы // ТИИЭР. 1976. Т.64. №3. С. 4–14.
10. *Хургин Я.И., Яковлев В.П.* Прогресс в Советском Союзе в области теории финитных функций и ее применений в физике и технике // ТИИЭР. 1977. Т.65. №7. С. 16–45.
11. *Джерри А.* Теорема отсчетов Шеннона, ее различные обобщения и ограничения // ТИИЭР. 1977. Т.65. №11. С. 53-89.
12. *Eldar Y.C.* Sampling Theory: Beyond Bandlimited Systems // University Printing House. Cambridge CB2 8BS. UK. 2015. 799 p.
13. *Птачек М.* Цифровое телевидение. Теория и техника // М.: Радио и связь. 1990. 528 с.
14. *Малла С.* Вейвлеты в обработке сигналов // М.: Мир. 2005. 672 с.
15. *Petersen D., Middleton D.* Sampling and Reconstruction of Wave-Number-Limited Functions in N-dimensional Euclidean Spaces // Information and Control. 1962. no. 5. pp. 279–323.
16. *Марчук Г.И., Агошков В.И.* Введение в проекционно-сеточные методы // М.: Наука. 1981. 416 с.
17. *Гребенников А.И.* Метод сплайнов и решение некорректных задач теории приближений // М.: Изд. МГУ. 1983. 208 с.
18. *Мирошниченко В.Л.* Об интерполяции и аппроксимации сплайнами // Вычислительные системы. 1983. Вып. 100. С. 83–100.
19. *Свиньин С.Ф.* Базисные сплайны в теории отсчетов сигналов // СПб: Наука. 2003. 118 с.
20. *Блаттер К.* Вейвлет-анализ. Основы теории // М.: Техносфера. 2006. 272 с.
21. *Ахмед Н, Рао К.* Ортогональные преобразования при обработке цифровых сигналов // М.: Связь. 1980. 248 с.
22. *Benedetto J., Ferreiro P.* Modern Sampling Theory: Mathematics and Applications // Springer Science and Business Media LLC. 2012. 158 p.
23. *Benedetto J.* Sampling Theory and Wavelets // Signal Processing for multimedia. Ed. J.S. Burns. IOS Press. 1999. pp.19–33.

References

1. Whittaker E. On the functions which are represented by expansions of the interpolation theory. *Proc. Roy. Soc. Edinburgh.* 1915. vol. 35. pp. 181–194.
2. Kotel'nikov V.A. [About bandwidth of “ether” and wire in telecommunications]. *Uspehi fizicheskih nauk – Advances in Physical Sciences.* 2006. vol. 176. no. 7. pp. 762–770. (In Russ.).
3. Shannon C. Communications in the presence of noise. *Proc. IRE.* 1949. vol. 37. no. 1. pp. 10–21.
4. Harkevich A.A. [About Kotelnikov's theorem]. *Radiotehnika – Radiotechnics.* 1958. vol. 1. no. 8. pp. 3–10. (In Russ.).
5. Papoulis A. [Error analysis in sampling theory]. *TIIER – Institute of engineers in electronics and radiotechnics proceedings.* 1966. vol. 54. no. 7. pp. 34-43. (In Russ.).
6. Hurgin Ja.I., Jakovlev V.P. *Finitnye funkcii v fizike i tehnike* [Finite functions in science and technics]. M.: Nauka. 1971. 408 p. (In Russ.).
7. Cybakov B.S., Jakovlev V.P. [About accuracy of recovery of functions with a finite number of terms of the Kotelnikov series] *Radiotehnika i jelektronika – Radiotechnics and electronics.* 1959. vol. 4. no. 3. pp. 542. (In Russ.).
8. Landau G. [Sampling method, information transfer and Nyquist frequency] *TIIER – Institute of engineers in electronics and radiotechnics proceedings.* 1967. vol. 55. no. 10. pp. 56–62. (In Russ.).
9. Slepjan A.D. [About bandwidth]. *TIIER – Institute of engineers in electronics and radiotechnics proceedings.* 1976. vol. 64. no. 3. pp. 4–14.
10. Hurgin Ja.I., Jakovlev V.P. [Progress in the Soviet Union in the field of finite functions theory and its applications in physics and technics] *TIIER – Institute of engineers in*

electronics and radiotechnics proceedings. 1977. vol. 65. no. 7. pp. 16-45. (In Russ.).
11. Jerri A. [Shannon's sampling theorem, it's various generalizations and constraint]. *TIIER – Institute of engineers in electronics and radiotechnics proceedings.* 1977. vol 65. no. 11. pp. 53–89. (In Russ.).
12. Eldar Y.C. Sampling Theory: Beyond Bandlimited Systems. University Printing House. Cambridge CB2 8BS. UK. 2015. 799 p.
13. Ptachek M. *Cifrovoe televidenie. Teorija i tehnika* [Digital television. Theory and Technics]. M.: Radio i svjaz', 1990. 528 p. (In Russ.).
14. Malla S. *Vejvlety v obrabotke signalov* [Wavelets in signal processing]. M.: Mir. 2005. 672 p. (In Russ.).
15. Petersen D., Middleton D. Sampling and Reconstruction of Wave-Number-Limited Functions in N-dimensional Euclidean Spaces. *Information and Control.* 1962. no. 5. pp. 279–323.
16. Marchuk G.I., Agoshkov V.I. *Vvedenie v proekcionno-setochnye metody* [Introduction to projection-grid methods]. M.: Nauka. 1981. 416 p. (In Russ.).
17. Grebennikov A.I. *Metod splajnov i reshenie nekorrektnyh zadach teorii priblizhenij* [Spline method and solutions of some incorrect problems in approximation theory]. M.: Izd. MGU. 1983. 208 p. (In Russ.).
18. Miroshnichenko V.L. [About spline interpolation and approximation] *Vychislitel'nye sistemy – Computer systems.* 1983. no. 100. pp. 83–100. (In Russ.).
19. Svin'in S.F. *Bazisnye splajny v teorii otschetov signalov* [Basic splines in the signals sampling theory]. SPb: Nauka. 2003. 118 p. (In Russ.).
20. Blatter K. *Vejvlet-analiz. Osnovy teorii* [Wavelet analysis: a primer]. M.: Tehnosfera. 2006. 272 p. (In Russ.).
21. Ahmed N., Rao K. *Ortogonal'nye preobrazovanija pri obrabotke cifrovyh signalov* [Orthogonal transforms for digital signal processing] M.: Svjaz'. 1980. 248 p. (In Russ.).
22. Benedetto J., Ferreiro P. Modern Sampling Theory: Mathematics and Applications. Springer Science and Business Media LLC. 2012. 158 p.
23. *Benedetto J.* Sampling Theory and Wavelets. Signal Processing for multimedia. Ed. J.S. Burns. IOS Press. 1999. pp.19–33.

Свиньин Сергей Федорович — д-р техн. наук, профессор, ведущий научный сотрудник лаборатории автоматизации научных исследований, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: цифровая обработка сигналов. Число научных публикаций — 160. svinyins@mail.ru; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7(812)323-5139, Факс: +7(812)328-4450.

Svinyin Sergey Fedorovich — Ph.D., Dr. Sci., associate professor, leading researcher of laboratory for research automation, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: digital processing of biomedical signals. The number of publications — 160. svinyins@mail.ru; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)323-5139, Fax: +7(812)328-4450.

Попов Александр Игоревич — к-т техн. наук, доцент кафедры прикладной информатики, Северный (Арктический) федеральный университет имени М.В. Ломоносова. Область научных интересов: цифровая обработка сигналов в электрофизиологии, информационные системы, автоматизация научных исследований. Число научных публикаций — 31. aleneus@gmail.com, http://dsplab.narfu.ru; Набережная Северной Двины, 17, Архангельск, 163000; р.т.: 8(8182)21-61-00, Факс: 8(8182)28-76-14.

Popov Aleksandr Igorevich — Ph.D., associate professor, associate professor of applied informatics department, Northern (Arctic) Federal University. Research interests: digital signal processing in electrophysiology, information systems, automation of researches. The number of publications — 31. aleneus@gmail.com, http://dsplab.narfu.ru; 17, Northern Dvina Embankment, Arkhangelsk, 163000, Russia; office phone: 8(8182)21-61-00, Fax: 8(8182)28-76-14.

РЕФЕРАТ

Свиньин С.Ф., Попов А.И. **Финитные базисные функции в задачах формирования выборок сигналов конечной протяженности.**

Теоремы Котельникова-Шеннона для выборок отсчетов непрерывных сигналов – функций времени с финитным спектром, базируются на применении кардинального ряда Уиттекера. Они предполагают бесконечную длительность ряда. В реальности имеет место его усечение, которое мало влияет на ошибку восстановления, если длительность сигнала значительно превышает (на несколько порядков) период самой низкочастотной его составляющей. Для функций времени это условие часто выполняется. Но сигналы конечной длительности, особенно с разрывными значениями на границах, не являются целыми функциями. Спектры таких сигналов инфинитны и их значения в области высоких частот должны учитываться.

Проблемы восстановления непрерывной информации по дискретным выборкам значительно возрастают, если пространства функций имеют размерности два, три и более, а формы носителей сигналов расширяются до прямоугольников, параллелепипедов, гиперкубов и т.д. Существуют методы многомерного анализа, ориентированные на принцип финитности спектров и на операции их периодического продолжения в области высоких частот. Такой подход требует рассмотрения периодических продолжений и для исходных сигналов, что порождает искажения, поскольку получающиеся многомерные картины далеко не всегда соответствует реальным физическим полям.

В статье оцениваются возможности и перспективы методов расчета необходимых выборок сигналов на основе теории финитных функций с конечной энергией с применением равенства Парсеваля. Исследуются соотношения между полной энергией в пространстве сигнала и последовательностью спектральных энергий коэффициентов разложения сигналов по базисным функциям с компактными носителями. В качестве примеров финитных базисных функций приводятся полиномиальные B-сплайны, а также ортонормированные вейвлеты семейства Добеши.

B-сплайнам соответствуют аналитические выражения для описания их спектров Фурье, которые инфинитны. Уровень энергии спектра используется для оценки шага выборки. Вейвлетам Добеши присущи алгоритмы быстрых вейвлет-преобразований. Их результаты в виде наборов коэффициентов, а также октавный принцип накопления спектральной энергии, с ростом числа итераций обеспечивают необходимую величину шага.

Предложенный энергетический критерий расчета длин выборок финитных сигналов с применением финитных базисов имеет значительные перспективы для расширения на область сигналов – функций двух, трех и более переменных. Этому способствует свойство сепарабельности многомерных B-сплайнов и ортонормированных вейвлет-функций с компактным носителем, а также подобное свойство их спектров.

SUMMARY

Svinyin S.F., Popov A.I. **Finite Basic Functions in the Tasks of Sampling Signals of Finite Qxtension.**

Kotelnikov-Shannon sampling theorems of continuous signals - time functions with finite spectrum are based on the application of the Whittaker cardinal series. They presuppose the infinite duration of the series. In reality, there is a truncation, which has insignificant effect on error recovery, if the duration of the signal significantly exceeds (by several orders of magnitude) the period of the low-frequency component of it. For a function of time, this condition is often performed. But the signals of finite duration, especially with discontinuous values at the borders, are not integral functions. The spectra of these signals are infinite and their values at high frequencies beyond the Nyquist frequency must be counted.

The problems regarding restoration of continuous information using discrete samples increase significantly, if the function spaces have two, three or more dimensions, and the forms of signal carriers are expanded to rectangles, parallelepipeds, hypercube, etc. The methods of multidimensional analysis, oriented to the principle of finite spectra and operations of their periodic continuation technique to the areas of high frequencies, exist. These methods require consideration of the periodic extension for the original signals, causing the distortions, because a multidimensional picture does not always correspond to the real physical field.

In the article, the possibilities and perspectives of methods of calculating the required step of signal samples based on the finite functions theory with finite energy and with the use of Parseval equality are estimated. We investigate the relation between the total energy of space signal and the sequence of spectral energy coefficients of decomposition for basic functions with compact carriers. As the examples of the finite basic functions, the polynomial B-splines and orthonormal wavelets Daubechies are used.

B-splines have analytical expression to describe their own Fourier spectra, which are infinite. The energy level of the spectrum is used to estimate a sampling step. As for Daubechies wavelets, they are characterized by fast algorithms of wavelet transforms. The results of these transforms obtained in the form of coefficients sequences, as well as the principle of octave spectral energy accumulation during increasing number of iterations, provide the necessary step.

The proposed energy criterion for computing the lengths of samples of finite signals with finite basic functions has considerable perspectives for expansion to spaces of functions with two, three, or more variables. This is achieved due to the separability property of multidimensional B-splines and orthonormal wavelet functions with compact carriers and similar property of their spectra.

УДК 519.246.87

Н.В. Абалов, В.В. Губарев

АВТОМАТИЧЕСКАЯ ГРУППИРОВКА КОМПОНЕНТ РАЗЛОЖЕНИЯ ВРЕМЕННОГО РЯДА ПРИ СИНГУЛЯРНОМ СПЕКТРАЛЬНОМ АНАЛИЗЕ

Абалов Н.В., Губарев В.В. **Автоматическая группировка компонент разложения временного ряда при сингулярном спектральном анализе.**

Аннотация. Сингулярный спектральный анализ (ССА) является сравнительно новым методом анализа временных рядов. ССА представляет особый интерес в приложении к анализу нестационарных, коротких и зашумлённых рядов. Одной из слабых сторон метода является то, что простые гармонические колебания, как и более сложные компоненты, анализируемого временного ряда раскладываются на более чем одну компоненту, что приводит к необходимости группировки связанных компонент для дальнейшего анализа. Данная проблема частично рассматривается в работе Александрова и Голяндиной (2005), преимущественно в приложении к проблеме идентификации чистых гармонических колебаний.

В данной работе предлагается более гибкий и обобщённый алгоритм для автоматической группировки компонент (а также его модификация), позволяющий группировать не только компоненты, соответствующие гармоническим колебаниям, но и компоненты, соответствующие амплитудно-модулированным колебаниям, затухающим колебаниям и др. Алгоритм был апробирован на искусственных наборах данных, содержащих в себе следующие распространенные формы компонент: гармоническое, амплитудно-модулированное и экспоненциально-затухающее колебания, сумма двух кривых Гаусса, а также их различные аддитивные комбинации. Экспериментально получены оценки качества группировки и показано, что показатели качества группировки у предложенных алгоритмов в среднем лучше на 26%, чем показатели известного алгоритма.

Ключевые слова: сингулярный спектральный анализ; ССА; временные ряды; группировка; идентификация.

Abalov N.V., Gubarev V.V. **Automatic Grouping of Time Series Decomposition Components in Singular Spectrum Analysis.**

Abstract. Singular spectrum analysis (SSA) is a relatively new method of time series analysis. SSA is of particular interest in application to analysis of non-stationary, short and noise time series. One of the drawbacks of SSA is that both simple harmonic oscillations and complex components of analyzed time series are decomposed into more than one component, which leads to the necessity of grouping related components for further analysis. This problem was partially addressed by Alexandrov, Golyandina (2005), mainly in application to the problem of identification of harmonic oscillations. In this paper, we present a more agile and generalized algorithm for automated grouping of components, which allows grouping not only harmonic oscillations, but also components corresponding to amplitude-modulated oscillations, fading oscillations and other. The algorithm was tested on synthetic time series, composed of common components: harmonic, amplitude-modulated, and exponentially damped oscillations, sum of two Gaussians, and their linear combinations. Experimental results of quality of grouping were obtained, showing that the proposed algorithm gives on average 26% better grouping results than an existing algorithm.

Keywords: singular spectrum analysis, SSA, time series, grouping, identification.

1. Введение.

Нестационарное временное поведение характерно для различных естественных, социально-экономических и технических

процессов. При обработке на цифровых вычислительных машинах они представляются эмпирическими данными, которые можно описать нестационарными временными рядами (ВР). Относительно новым и набирающим распространение методом анализа таких ВР является метод сингулярного спектрального анализа (ССА). Его основой послужили методы главных компонент и теории динамических систем. Описание метода и ссылки на ключевые работы в этой области можно найти в [1–6].

Среди основных сильных сторон ССА можно отметить (см., например, [1, 3]) то, что он не предъявляет строгих требований к стационарности ряда, применим к зашумленным и коротким рядам, позволяет выделять как периодические, так и сложные нестационарные компоненты. Слабой стороной метода является то, что он не дает компактного аналитического модельного представления ряда и требует значительного объема ручной работы в диалоговом режиме.

Ранее (см., например, [7, 8]) нами был предложен метод автоматической идентификации нестационарных временны рядов на основе вариативного моделирования, объединяющий методы ССА и моделетеки. Одной из проблем, которая возникает при реализации этого метода, является автоматическая группировка компонент ССА разложения, относящихся к разным составляющим ряда. Она возникает из-за того, что метод ССА в общем случае раскладывает гармонические и более сложные, например затухающие, колебания более чем на одну компоненту. При этом на практике крайне желательно, чтобы компоненты разложения, соответствующие одной составляющей модельного представления ВР (такой как гармоническое и экспоненциально затухающее колебание, амплитудно-модулированное колебание или вейвлет и т.п.), были сгруппированы, позволяли компактно представить и воспроизвести связанную с ними составляющую исходного ряда, а также использовать их при дальнейшем анализе, исследовании ВР. Желательно чтобы такая группировка позволила значительно сократить время автоматического поиска и время подстройки моделей базовых компонент по типу того, как это делается в методе моделетеки, а также упростить результирующую модель.

Проблема группировки компонент возникает не только при совместном использовании ССА и моделетеки, но и при применении лишь ССА, так как исследователь должен вручную выбрать интересующие его компоненты и сгруппировать их. Зачастую при использовании ССА применяют укрупненную группировку, относя компоненты к одной из трех групп: тренд, колебания, шум. При таком крупном группировании, во-первых, понижается наглядность и информатив-

ность получаемых групп, во-вторых, сохраняется проблема компактности аналитического представления группы компонент, что ухудшает интерпретируемость и дальнейшее исследование результатов ССА.

Таким образом, возникает необходимость в алгоритме автоматической группировки связанных компонент разложения ССА. Цель работы – разработка алгоритма группирования компонент, позволяющего автоматизировать данный этап ССА и сократить объем вычислений в задачах вариативной идентификации [7]. В данной работе представлен разработанный алгоритм и проводится его сравнение с существующим алгоритмом.

2. Методы.

2.1. Сингулярный спектральный анализ. Для лучшего понимания предлагаемого метода, кратко рассмотрим на примере одномерного ряда Y_N основные этапы ССА, позволяющего разложить исходный временной ряд $Y_N = (y_0, y_1, \dots, y_{N-1})$ длины N, где $y_i = y(i\Delta t)$, $i = 1, \dots, N-1 = \overline{1, N-1}$, на набор аддитивных компонент. В ССА можно выделить два укрупненных этапа: разложение и восстановление.

*Разложение***.** Первым шагом этапа является преобразование (вложение) одномерного временного ряда Y_N в траекторную матрицу $\mathbf{X}$ размером $L \times K$, где L – длина скользящего вдоль ВР окна (гусеницы), $1 < L < N$, $K = N - L + 1$, $\mathbf{X} = [X_1, X_2, \dots, X_K]$, $X_h = (y_{h-1}, \dots, y_{h+L-2})^T$, $h = 1, \dots, K$. Следующий шаг – вычисление матрицы $\mathbf{X}\mathbf{X}^T$ и ее разложение по собственным векторам. Результатом шага является набор собственных троек $\left(\sqrt{\lambda_j}, U_j, V_j\right)$, $j = 1, \dots, d$, упорядоченных по убыванию ненулевых собственных чисел λ_j, где U_j – собственный вектор, а $V_j = \sqrt{\lambda_j}\mathbf{X}^T U_j$ – факторный вектор.

*Восстановление***.** На данном этапе производится группировка компонент в соответствии с интересами исследователя и восстановление сгруппированных компонент и всего ряда (численная замена исходного ряда новым, полученным суммированием значений отобранных и сгруппированных собственных (восстановленных) компонент). Для получения восстановленной компоненты RC_j, соответствующей одной j-ой тройке, необходимо вычислить диагональное усреднение (ганкелизацию) матрицы $\sqrt{\lambda_j}U_j{V_j}^T$. Рассматриваемые далее алгоритмы применяются после этапа разложения, когда получены все тройки $\left(\sqrt{\lambda_j}, U_j, V_j\right)$, а также, в отдельных случаях, вычислены соответствующие этим тройкам восстановленные компоненты RC_j.

2.2. Алгоритм группировки Ф. И. Александрова, Н. Э. Голяндиной. В работе [9] при решении задачи идентификации различных компонент (трендовой, колебательной или шумовой) косвенно,

как этап выполнения идентификации, решается задача группировки только пар компонент, соответствующих чистым гармоникам.

Пусть Π_Z – значение нормированной периодограммы заданного ряда Z, имеющего длину L. $\Pi_Z(k)$ – значение периодограммы, соответствующее частоте k/L, где $0 \le k \le L/2$ – индекс частоты. Суть предложенного Александровым и Голяндиной алгоритма можно кратко изложить в виде последовательности следующих действий:

1. Перебираем пары последовательных собственных троек $\left(\sqrt{\lambda_j}, U_j, V_j\right), \left(\sqrt{\lambda_{j+1}}, U_{j+1}, V_{j+1}\right), j = 1, \dots, d-1.$
2. Для каждой пары вычисляем показатель:

$$\rho_{j,j+1} = \frac{1}{2}\max_{0 \le k \le L/2}\left(\Pi_{U_j}(k) + \Pi_{U_{j+1}}(k)\right). \quad (1)$$

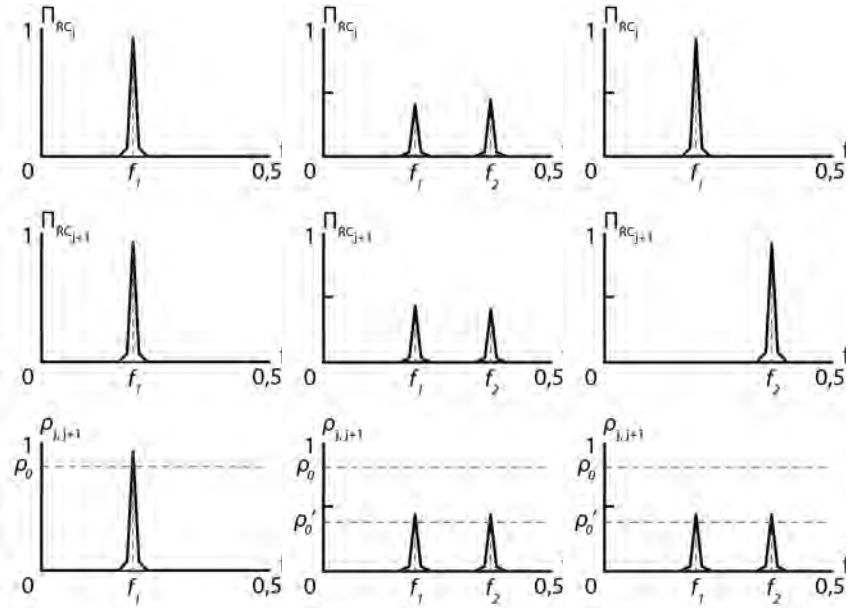
3. Если для текущего j значение $\rho_{j,j+1} \ge \rho_0$, где $\rho_0 \in [0,1]$ заданный пользователем предел, то тройки $\left(\sqrt{\lambda_j}, U_j, V_j\right)$ и $\left(\sqrt{\lambda_{j+1}}, U_{j+1}, V_{j+1}\right)$ считаются относящимися к одной компоненте (гармонике) и группируются, в противном случае – не относятся.
4. Переходим к следующему j, т.е. следующим нерассмотренным парам собственных троек, пока не переберем их все.

Согласно данному алгоритму, пара соседних $(j, j+1)$ собственных троек идентифицируется как относящаяся к одному гармоническому колебанию (пара группируется и восстанавливается как одна гармоника – составляющая ряда) при условии, что пики периодограмм Π_{U_j}, $\Pi_{U_{j+1}}$, вычисленных по собственным векторам U_j, U_{j+1} соответственно, приходятся на одинаковые частоты.

Отметим, что в (1) вместо периодограмм Π_{U_j} и $\Pi_{U_{j+1}}$, вычисленных по собственным векторам, можно использовать значения периодограмм Π_{RC_j} и $\Pi_{RC_{j+1}}$, вычисленных по значениям восстановленных компонент RC_j и RC_{j+1} (при условии, что каждая восстановленная компонента RC_j соответствует одной j-ой собственной тройке).

Указанный алгоритм достаточно жесток, причем чем больше ρ_0, тем жестче условие. При больших ρ_0, это затрудняет применение алгоритма в условиях зашумленности ряда или наличия сложных нестационарных компонент. При низких же значениях ρ_0 возможны ложные выводы и некорректные результаты. Алгоритм, фактически, не предусматривает случаев группировки компонент, являющихся составными для амплитудно-модулированных и затухающих колебаний (поскольку авторами [9] изначально ставилась задача идентификации чистых гармонических колебаний). В дальнейшем будем обозначать этот алгоритм как HG (harmonic grouping).

Рассмотрим пример, как жесткость алгоритма может приводить к низкому качеству группировки. Положим, что $\rho_0 = 0{,}8$. Рассмотрим рисунок 1.



а) *б)* *в)*

Рис. 1. Примеры группировки компонент

Случай *а)* является наиболее простым, именно для него предлагался алгоритм в [9]. Поскольку пики двух периодограмм приходятся на одну частоту, совпадающую с f_1, каждая нормированная периодограмма содержит только один пик (их значения, как следствие, близки к 1), то оценка $\rho_{j,j+1}$ имеет значение близкое к единице, превышающие ρ_0. В случае *б)*, поскольку каждая нормированная периодограмма содержит по два пика, то максимальное значение нормированной периодограммы не превышает 0,5. Тогда при вычислении показателя $\rho_{j,j+1}$ по (1) его значение также не может превышать 0,5, т.е. будет меньше, чем ρ_0. Как следствие данная пара, соответствующая одной компоненте, не будет автоматически сгруппирована. Чтобы автоматически группировались и такие компоненты, необходимо понизить ρ_0. Для группировки данной пары необходимо снизить значение ρ_0 и принять его равным, например, $\rho_0' = 0{,}4$. Тогда возникает ситуация *в)*, когда рассматривается пара, которая не должна быть сгруппирована, по-

скольку относится к двум разным чистым гармоникам с частотами f_1 и f_2 соотвественно. Поскольку каждая из $j, j+1$ периодограмм содержит только один пик, максимальное значение данных пиков будет равно 1 или близко к нему. Тогда $\rho_{j,j+1}$ в соответствии с (1) будет равно или незначительно меньше 0,5, т.е. больше, чем выбранное нами из-за случая *б)*, $\rho'_0 = 0{,}4$ и две разные гармоники будут ошибочно автоматически сгруппированы в одну компоненту.

2.3. Предлагаемый алгоритм. Ниже описывается более гибкий алгоритм, направленный на группировку собственных компонент, относящихся к таким элементарным составляющим исходного ВР как: гармонические, амплитудно-модулированные, экспоненциально затухающие колебания и т.п. В дальнейшем будем обозначать этот алгоритм как GG (generalized grouping).

Алгоритм может быть записан в виде следующей последовательности действий:

1. В отличие от HG, для всех компонент рассчитываем матрицу близости собственных чисел $\Delta = [\delta_{ij}]$, $i, j = 1, \dots, d$, где

$$\delta_{ij} = \frac{\min\limits_{\lambda}(\lambda_i, \lambda_j)}{\max\limits_{\lambda}(\lambda_i, \lambda_j)}. \qquad (2)$$

2. Рассчитываем матрицу «смежности» компонент $\mathbf{G} = [g_{ij}]$, $i, j = 1, \dots, d$,

$$g_{ij} = \begin{cases} c_{ij}, & \delta_{ij} \geq \rho_1, \\ 0, & \text{иначе}, \end{cases}$$

где c_{ij} – бинарный показатель близости двух компонент (далее в работе будут рассмотрены два таких показателя).

Итогом выполнения операций 1–2 является матрица группировки $\mathbf{G}$ (аналогичная матрице смежности в теории графов), содержащая 1 в ячейках i, j, если пара компонент i, j принадлежат одной группе, иначе 0.

3. Объединяем в одну группу составляющей ряда те собственные компоненты, для которых $g_{ij} = 1$.

Первый показатель близости c'_{ij} основан на использовании степени коэффициента связи между значениями периодограмм восстановленных компонент (например, коэффициента корреляции Пирсона, Спирмена, конкоров Губарева, корреляционных отношений и т.п.).

Здесь $c'_{ij} = 1$, тогда и только тогда, когда $corr(RC_i, RC_j) \geq \rho_c$, $\rho_c \in [0,1]$.

Использование алгоритма GG с показателем в виде коэффициента корреляции Пирсона на рассматриваемых примерах, как будет показано ниже, в среднем дает лучшие результаты, чем алгоритм HG, особенно в случае наличия амплитудно-модулированных и других компонент, отличных от чистой гармоники. Чем выше значение ρ_c, тем жестче условие.

Второй показатель близости c''_{ij} основан на использовании «гибкого» сравнения множеств частот, соответствующих максимальным значениям периодограмм. Введем следующие обозначения:

$K_{RC_j} = \max_{o \le k \le L/2}\left(\Pi_{RC_j}(k)\right)$ – максимальное значение периодограммы восстановленной компоненты RC_j;

$$F_{RC_j} = \left\{k | \Pi_{RC_j}(k) \ge \rho_p K_{RC_j}\right\}, j = 1, \dots, d, \quad (3)$$

F_{RC_j} – упорядоченное по возрастанию значений множество из n индексов k частот, соответствующих первым n максимальным значениям периодограммы Π_{RC_j}, превышающим порог $\rho_p K_{RC_j}$, где ρ_p – задаваемая величина порога.

Величина порогового значения ρ_p ($\rho_p \in [0,1]$) позволяет регулировать исключение низких значений периодограмм. Например, в случае, когда периодограмма содержит лишь один «доминирующий» пик и остальные будут ниже порогового значения (для чистого гармонического колебания), независимо от n будет выбрана лишь одна частота. Значение $n = 2$ – является достаточным для выделения двух пиковых значений периодограммы, позволяющих идентифицировать амплитудно-модулированное колебание. Чем выше значение ρ_p, тем жестче условия на отбор «значимых» частот.

Здесь $c''_{ij} = 1$ тогда и только тогда, когда $\forall h = 1, \dots, m$ имеем $\frac{\left|F_{RC_i}(h) - F_{RC_j}(h)\right|}{L/2} \le \rho_2$, где $|\ |$ – модуль, $F_{RC_i}(h)$ – h-ое значение индекса частоты из множества F_{RC_i}, ρ_2 – порог предельного расхождения индексов частот, $\rho_2 \in [0,1]$, $m = \min\left(\#F_{RC_i}, \#F_{RC_j}\right)$, где # – обозначает мощность соответствующих множеств. Чем меньше значение ρ_2, тем жестче условие.

Использование данного показателя c''_{ij} должно позволить учесть размытие и слабое смещение («дрифт») пиковых значений периодограмм при повышении уровня соотношения сигнала к шуму.

Рассмотрим основные отличия предложенного алгоритма, являющегося расширением идей алгоритма Александрова и Голяндиной, от существующего алгоритма. Он является расширением алгоритма Александрова и Голяндиной. Они заключаются в следующем:

а) Вместо строго последовательного попарного рассмотрения еще не сгруппированных компонент предлагается рассматривать все комбинации пар компонент, имеющих близкие значения собственных числе в соответствии с оценкой (2).

б) Оценка близости частотного состава осуществляется не по совпадению значений частот, соответствующих только одному доминирующему пику в периодограммах собственных векторов, а по одному из двух предложенных критериев, первый из которых позволяет учесть весь частотный состав, второй – более одного пика в периодограмме и слабое смещение («дрифт») значений периодограммы из-за высокого уровня шума.

Первая особенность позволяет обобщить правило группирования и обеспечить большую гибкость алгоритма в условиях сильной зашумленности ряда или наличия сложных нестационарных компонент. Вторая особенность позволяет сместить фокус с группировки только чистых гармонических колебаний на группировку амплитудно-модулированных гармонических колебаний (периодограммы которых имеют пики на двух частотах), затухающих колебаний и других сложных по форме компонент.

3. Методология тестирования и сравнения алгоритмов. Для тестирования алгоритмов используется метод идеального сигнала. Многократно повторялись опыты по группировке собственных компонент для различных значений соотношения уровня сигнала к шуму $s \in S$ и различных наборов составляющих исходного ряда C^h, $h \in H$. Для оценки качества группировки используется следующий подход.

Для каждой пары искусственного набора компонент и заданного соотношения уровня сигнала к шуму $\langle h, s\rangle \in H \times S$, $H = \{0,1, \dots ,5\}$, $S = \{0{,}05; 0{,}1; 0{,}25; 0{,}5\}$:

1. Повторить (200 раз) следующие шаги:

1.1. Генерируется искусственный временной ряд $x(t) = \sum_{i=1}^{N_h} c_i^h(t) + \varepsilon(t)$, где $N_h = \#C^h$ – количество составляющих (искусст-

венно заданных компонент) в h-ом наборе компонент, $c_i^h(t)$ – случайная реализация i-ой компоненты h-ого набора компонент C^h, $\varepsilon(t)$ – случайный шум, имеющий нормальное распределение с нулевым средним и стандартным отклонением $\sigma_\varepsilon = s \cdot \sigma_\Sigma$, где σ_Σ – стандартное отклонение $\sum_{i=1}^{N_h} c_i^h(t)$.

1.2. Производится автоматическая группировка собственных компонент разложения $x(t)$ рассматриваемыми алгоритмами группировки.

1.3. Заложенные в искусственный имитируемый ряд составляющие сопоставляются с компонентами, полученными в результате группировки. Для этого для каждой составляющей $c_i^h(t)$, заложенной в ряд, находится такая собственная компонента из результата группировки $g_j^h(t)$, которая обеспечивает наибольше значение коэффициента детерминации $R^2{}_i = R^2\left(c_i^h(t), g_j^h(t)\right)$.

1.4. Вычисляется среднее по всем N_h компонентам значение $\overline{R^2}$ полученных для каждого алгоритма коэффициентов детерминации $R^2{}_i$, $i = 1, \dots, N_h$.

2. Для каждого опыта и алгоритма вычисляются следующие оценки показателя качества группировки: $\mu_{\overline{R^2}}$ – среднее значение $\overline{R^2}$ и $\sigma_{\overline{R^2}}$ – СКО $\overline{R^2}$, полученные по 200 повторениям опыта.

Коэффициент детерминации R^2 для каждой составляющей (искусственной компоненты) отражает долю её дисперсии, объясняемой рассматриваемой моделью (выделенной группой собственных компонент).

Пусть X~$U(a; b)$ обозначает, что X – случайная переменная, непрерывно равномерно распределенная на открытом интервале $(a; b)$. Для всех компонент $x = 1, \dots, len$. Если не оговорено иное, то φ_i~$U(-\frac{\pi}{2}; \frac{\pi}{2})$, $len = \lfloor L \rfloor$, где L~$U(182; 730)$. Случайные переменные остаются постоянными для одного опыта (повторения). Новое значение переменной выбирается случайно в соответствии с ограничениями при каждом новом опыте. Описание случайных компонент $c_i^h(t)$, соответствующих тестовому набору C^h, приведено в таблице 1.

Напомним обозначения: HG – алгоритм Александрова и Голяндиной; GG1 – предложенный в работе алгоритм, использующий первый показатель; GG2 – предложенный в работе алгоритм, использующий второй показатель.

Таблица 1. Описание тестовых наборов

k	Аналитическая запись компоненты	Примечания
0	$c_1(t) = \sin(2\pi t/T_1 + \varphi_1)$, $T_1 \sim U(5; 5 + \frac{1}{2} len)$;	Простое гармоническое колебание;
1	$c_1(t) = \sin(2\pi t/T_1 + \varphi_1)\sin(2\pi t/T_2 + \varphi_2)$, $T_1 \sim U(5; 5 + \frac{1}{3} len)$, $T_2 \sim U(\frac{1}{3} len; \frac{2}{3} len)$;	Амплитудно-модулированное колебание;
2	$c_1(t) = \sin(2\pi t/T_1 + \varphi_1)\, e^{-\gamma t/len}$, $T_1 \sim U(5; 5 + \frac{1}{2} len)$; $\gamma \sim U(\frac{1}{2}; 5)$;	Затухающее колебание;
3	$c_1(t) = a_1 e^{-\left(\frac{t-b_1}{c_1}\right)^2} + a_2 e^{-\left(\frac{t-b_2}{c_2}\right)^2}$, $a_1, a_2 \sim \pm U(\frac{1}{2}; 1)$, $b_1, b_2 \sim U(0; \frac{1}{2} len)$, $c_1, c_2 \sim U(0; \frac{1}{3} len)$;	Сумма двух гауссовых кривых;
4	$c_1(t) = a_1 \sin(2\pi t/T_1 + \varphi_1)$, $a_1 \sim U(7; 9)$, $T_1 \sim U(5; 10)$;	Сумма нескольких гармонических колебаний;
	$c_2(t) = a_2 \sin(2\pi t/T_2 + \varphi_2)$, $a_2 \sim U(1; 4)$, $T_2 \sim U(28; 32)$;	
	$c_3(t) = a_3 \sin(2\pi t/T_3 + \varphi_3)$, $a_3 \sim U(10; 18)$, $T_2 \sim U(13; 15)$;	
	$c_4(t) = a_4 \sin(2\pi t/T_4 + \varphi_4)$, $a_4 \sim U(15; 21)$, $T_2 \sim U(58; 62)$;	
5	$c_1(t) = a_1 \sin\left(2\pi t/T_1 + \frac{1}{2}\right)$, $a_1 \sim U(\frac{9}{2}; \frac{13}{2})$, $T_1 \sim U(2 \cdot 365; 3 \cdot 365)$;	Ряд, содержащий все рассмотренные выше компоненты; $len = 2 \cdot 365$.
	$c_2(t) = a_2 \sin(2\pi t/T_2 + 1)$, $a_2 \sim U(\frac{3}{2}; \frac{5}{2})$, $T_2 \sim U(5; 7)$;	
	$c_3(t) = -a_3 \sin(2\pi t/T_{31} + 1)\sin(2\pi t/T_{32})$, $a_3 \sim U(\frac{7}{2}; \frac{9}{2})$, $T_{31} \sim U(10; 12)$, $T_{32} \sim U(150; 170)$;	
	$c_4(t) = -a_4 \sin(2\pi t/T_4 + 1)$, $a_4 \sim U(\frac{1}{2}; \frac{3}{2})$, $T_4 \sim U(29; 31)$;	
	$c_5(t) = a_{51}\exp\left(-\left(\frac{t-b_{51}}{c_{51}}\right)^2\right) + a_{52}\exp\left(-\left(\frac{t-b_{52}}{c_{52}}\right)^2\right)$, $a_{51}, a_{52} \sim U(\frac{1}{2}; \frac{3}{2})$, $b_{51}, b_{52} \sim U(0; \frac{1}{2} len)$, $c_{51}, c_{52} \sim U(0; \frac{1}{3} len)$;	

Выбор параметров осуществлялся на основе предварительных экспериментов и рекомендаций из [9] для алгоритма HG. Опыты проводились при следующих значениях параметров. Для алгоритма HG: $\rho_0 = 0{,}8$. Для алгоритма GG1 и GG2 $\rho_1 = 0{,}8$, $\rho_c = 0{,}8$, $\rho_p = 0{,}8$, $\rho_2 = 0{,}05$.

Таблица 2. Результаты экспериментов

h	s	HG $\mu_{\overline{R^2}}$	HG $\sigma_{\overline{R^2}}$	GG1 $\mu_{\overline{R^2}}$	GG1 $\sigma_{\overline{R^2}}$	GG2 $\mu_{\overline{R^2}}$	GG2 $\sigma_{\overline{R^2}}$
0	0,05	0,84	0,11	0,98	0,06	1,00	0,00
	0,1	0,82	0,10	0,98	0,06	1,00	0,00
	0,25	0,83	0,12	0,99	0,06	1,00	0,00
	0,5	0,83	0,11	0,98	0,06	1,00	0,00
1	0,05	0,50	0,07	0,76	0,16	1,00	0,01
	0,1	0,49	0,10	0,71	0,20	1,00	0,03
	0,25	0,50	0,08	0,73	0,16	0,99	0,04
	0,5	0,50	0,07	0,67	0,14	0,99	0,02
2	0,05	0,81	0,09	0,93	0,12	1,00	0,00
	0,1	0,80	0,09	0,93	0,12	1,00	0,00
	0,25	0,80	0,09	0,94	0,12	1,00	0,01
	0,5	0,81	0,08	0,96	0,10	1,00	0,01
3	0,05	0,70	0,20	0,73	0,19	0,81	0,15
	0,1	0,71	0,21	0,75	0,21	0,83	0,15
	0,25	0,69	0,21	0,73	0,20	0,81	0,17
	0,5	0,70	0,21	0,73	0,21	0,82	0,15
4	0,05	0,86	0,07	0,99	0,02	0,95	0,14
	0,1	0,87	0,06	0,99	0,02	0,96	0,14
	0,25	0,84	0,07	0,95	0,08	0,86	0,18
	0,5	0,75	0,12	0,81	0,18	0,70	0,21
5	0,05	0,64	0,04	0,77	0,04	0,79	0,03
	0,1	0,66	0,04	0,77	0,04	0,79	0,02
	0,25	0,64	0,04	0,74	0,03	0,78	0,02
	0,5	0,62	0,05	0,69	0,05	0,73	0,05
Среднее:		0,72	0,10	0,84	0,11	0,91	0,06

$\mu_{\overline{R^2}}$ – усредненное по 200 повторениям опыта значение оценок качества группировки для заданного набора данных и уровня шума.

$\sigma_{\overline{R^2}}$ – СКО оценок качества группировки для заданного набора данных и уровня шума.

4. Результаты и их обсуждение. Полученные результаты экспериментов представлены в 2. Как видно из таблицы 2, предложенный алгоритм в большинстве случаев показал лучшие результаты: усредненная по многократным повторениям оценка качества группировки $\mu_{\overline{R^2}}$ (средняя оценка $\overline{R^2}$) для предлагаемого алгоритма в большинстве опытов превышает аналогичную для существующего алгоритма. Особо это заметно в случае опытов с амплитудно-модулированной компонентой. Поскольку в данном случае значение $\rho_{j,j+1}$ в (1) мало, т.к. значения периодограммы почти равномерно распределяется между двумя пиками и получаемое значение зачастую значительно меньше значения порога ρ_0,

значительное снижение порогового значения может привести к ошибкам группировки. Кроме того среднее значение СКО оценки качества группировки для алгоритма GG2 ниже такового у алгоритма HG в среднем на 32%, что позволяет предположить, что предлагаемый алгоритм более устойчив к статистическим погрешностям.

5. Заключение. В работе был предложен алгоритм автоматической группировки компонент разложения для метода ССА и его две модификации GG1 и GG2. Алгоритм был апробирован на искусственных данных. Предложенный алгоритм был сравнен по качеству группировки с существующим алгоритмом. Алгоритмы GG1 и GG2 показали лучшее качество группировки (R^2 для GG1 в среднем на 16,67% выше, а для GG2 – на 26,39% выше, чем для существующего алгоритма). Это особо заметно в случаях, когда временной ряд содержит амплитудно-модулированные колебания.

Одним из направлений дальнейшей работы по исследованию алгоритма является выработка рекомендаций по выбору значений пороговых параметров в зависимости от характера составляющих ряда и уровня шума, экспериментальная проверка алгоритмов на примере других составляющих ряда.

Литература

1. *Данилов Д.Л., Жиглявский А.А.* Главные компоненты временных рядов: метод Гусеница // СПб: Издательство Санкт-Петербургского университета. 1997. 307 с.
2. Time series analysis and forecasting, Caterpillar SSA method. URL: http://www.gistatgroup.com/ (дата обращения: 10.04.2014).
3. *Vautard R., Yiou P., Ghil M.* Singular-spectrum analysis: A toolkit for short, noisy chaotic signals // Phys. D Nonlinear Phenom. Elsevier. 1992. vol. 58. no. 1-4. pp. 95–126.
4. *Hassani H.* Singular Spectrum Analysis: Methodology and Comparison // J. Data Sci. University Library of Munich. Germany. 2007. vol. 5. no. 4991. pp. 239–257.
5. *Hassani H., Thomakos D.* A review on singular spectrum analysis for economic and financial time series // Stat. Interface. 2010. vol. 3. pp. 377–397.
6. *Ghil M., Taricco C.* Advanced spectral analysis methods // In Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models. 1997. pp. 137–159.
7. *Абалов Н.В., Губарев В.В., Альсова О.К.*. Использование методов сингулярного спектрального анализа и моделетеки при идентификации временных рядов // Труды СПИИРАН. 2014. Вып. 35. С. 49–63.
8. *Gubarev V.V., Alsova O.K., Abalov N.V., Melnikov G.A.* Use of variative modeling for the identification of random signals // Proc. of 7th International Forum on Strategic Technology (IFOST). Tomsk. 2012. vol. 1. pp. 739–742.
9. *Alexandrov Th., Golyandina N.* Automatic extraction and forecast of time series cyclic components within the framework of SSA // Proc. 5th St. Petersburg. Work. Simulation. 2005. pp. 45–50.

References

1. Danilov D., Zhigljavsky A.A. *Glavnye komponenty vremennyh rjadov: metod Gusenica* [Principal Components of Time Series: the Caterpillar Method]. SPB: St. Petersburg University. 1997. 307 p. (In Russ.).

2. Time series analysis and forecasting. Caterpillar SSA method. Available at: http://www.gistatgroup.com/ (accessed: 10.04.2014).
3. Vautard R., Yiou P., Ghil M. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Phys. D Nonlinear Phenom.* Elsevier. 1992. vol. 58. no. 1-4. pp. 95–126.
4. Hassani H. Singular Spectrum Analysis: Methodology and Comparison. *J. Data Sci.* University Library of Munich, Germany. 2007. vol. 5. no. 4991. pp. 239–257.
5. Hassani H., Thomakos D. A review on singular spectrum analysis for economic and financial time series. *Stat. Interface*. 2010. vol. 3. pp. 377–397.
6. Ghil M., Taricco C. Advanced spectral analysis methods. *In Past and Present Variability of the Solar-Terrestrial System: Measurement, Data Analysis and Theoretical Models.* 1997. pp. 137–159.
7. Abalov N.V., Gubarev V.V., Alsova O.C. [Use of Methods of Singular Spectral Analysis and Modeleteka for the Identification of Time Series]. *Trudy SPIIRAN – SPIIRAS Proceedings.* 2014. vol. 35. pp. 49–63. (In Russ.).
8. Gubarev V.V., Alsova O.K., Abalov N.V., Melnikov G.A. Use of variative modeling for the identification of random signals. Proc. of 7th International Forum on Strategic Technology (IFOST). Tomsk. 2012. vol. 1. pp. 739–742.
9. Alexandrov Th., Golyandina N. Automatic extraction and forecast of time series cyclic components within the framework of SSA. Proc. 5th St. Petersburg. Work. Simulation. 2005. pp. 45–50.

Абалов Николай Владимирович — аспирант кафедры вычислительной техники, ФГБОУ ВПО Новосибирский государственный технический университет. Область научных интересов: интеллектуальный анализ данных, вариативное моделирование. Число научных публикаций — 5. nickabalov@yahoo.com; пр. К. Маркса, 20, Новосибирск, 630073; р.т.: +7-913-714-97-03.

Abalov Nikolay Vladimirovich — Ph.D student of computer sciences department, Novosibirsk State Technical University (NSTU). Research interests: intellectual data analysis, variative modeling. The number of publications — 5. nickabalov@yahoo.com; 20, Prospekt K. Marksa, Novosibirsk, 630073; office phone: +7-913-714-97-03.

Губарев Василий Васильевич — д-р техн. наук, профессор, заслуженный деятель науки Российской Федерации, заслуженный работник высшей школы Российской Федерации, профессор кафедры вычислительной техники, ФГБОУ ВПО Новосибирский государственный технический университет (НГТУ). Область научных интересов: идентификация, измерение характеристик, имитация и прогнозирование случайных сигналов; вероятностное моделирование реальных объектов; статистические прикладные информационные системы; системный анализ в экспериментальных исследованиях; интеллектуальный анализ данных и вариативное моделирование; концептуальные основы информатики. Число научных публикаций — 500. gubarev@vt.cs.nstu.ru; пр. К. Маркса, 20, Новосибирск, 630073; р.т.: +7(383)346-11-33.

Gubarev Vasiliy Vasilyevich — Ph.D., Dr. Sci., professor, honored scientist of Russian Federation, honored worker of higher school of Russian Federation, professor of computer sciences department, Novosibirsk State Technical University (NSTU). Research interests: identification, measurement of characteristics, simulation and prediction of random signals; probabilistic modeling of real objects; applied statistical information systems; system analysis in experimental research; intellectual data analysis and variative modeling; conceptual foundations of informatics. The number of publications — 500. gubarev@vt.cs.nstu.ru; 20, Prospekt K. Marksa, Novosibirsk, 630073, Russia; office phone: +7(383)346-11-33.

РЕФЕРАТ

Абалов Н.В., Губарев В.В. **Автоматическая группировка компонент разложения временного ряда при сингулярном спектральном анализе.**

Статья посвящена рассмотрению проблемы автоматической группировки компонент разложения при сингулярном спектральном анализе (ССА). В работе предложен алгоритм для автоматической группировки компонент при ССА. Приведены результаты его апробации на искусственных данных и сравнения с существующим алгоритмом.

ССА является сравнительно новым методом анализа временных рядов. ССА представляет особый интерес в приложении к анализу нестационарных, коротких и зашумлённых рядов. Одной из слабых сторон метода является то, что простые гармонические колебания, как и более сложные компоненты, анализируемого временного ряда раскладываются на более чем одну компоненту, что приводит к необходимости группировки связанных компонент для дальнейшего анализа.

В работе рассматривается существующий алгоритм группировки гармонических компонент, предложенный Ф. И. Александровым, Н. Э. Голяндиной в приложении к задаче идентификации тренда и чистых гармонических колебаний во временных рядах. На ряде примеров показано, что существующий алгоритм жесток и малопригоден для решения задачи группировки сложных компонент нестационарных временных рядов.

Предложен алгоритм, направленный на группировку собственных компонент, относящихся к таким составляющим исходного временного ряда как: гармонические, амплитудно-модулированные, экспоненциально затухающие колебания и т.п.

Приведены результаты апробирования предложенного алгоритма на искусственных наборах данных. Экспериментально получены оценки качества группировки и показано, что показатели качества группировки у предложенных алгоритмов в среднем лучше на 26%, чем показатели известного алгоритма.

SUMMARY

Abalov N.V., Gubarev V.V. **Automatic Grouping of Time Series Decomposition Components in Singular Spectrum Analysis.**

The paper discusses the problem of automated grouping of decomposition components in singular spectrum analysis (SSA). In the paper, a new algorithm for automated grouping of decomposition comments in SSA is presented. The results of its approbation on synthetic time series and comparison to existing algorithm are presented.

SSA is a relatively new method of time series analysis. SSA is of great interest in application to analysis of non-stationary, short and noisy time series. One of the drawback of SSA is the fact that simple harmonic components and complex components of analyzed time series are decomposed into more than one component, which leads to a need for a grouping of such related comments for further analysis.

In the paper, an existing algorithm of grouping, proposed by Alexandrov Th., Golyandina N. in application to identification of trend and pure harmonic components, is considered. Several examples are provided to show that this algorithm is strict and might be unsuitable for solving the problem of grouping of complex components in non-stationary time series.

An algorithm is proposed for automated grouping of such components as harmonic, amplitude-modulated, and exponentially damped oscillations, etc.

Results of approbation of the algorithm on synthetic data are provided. Experimental results of quality of grouping were obtained, showing that the proposed algorithm gives on average 26% better grouping results than an existing algorithm.

УДК 006.72

А.А. Воевода, Д.О. Романников

АЛГОРИТМ ОБЪЕДИНЕНИЯ ЧАСТЕЙ ОРИЕНТИРОВАННОГО ГРАФА

Воевода А.А., Романников Д.О. **Алгоритм объединения частей ориентированного графа.**

Аннотация. Рассматривается задача объединения графов с общей частью, которые были получены в результате серии моделирований сети Петри с использованием программного пакета Colored Petri Nets Tools, в котором адресное пространство процесса ограничено 2^{32} байтами, начиная с различных вершин и при различных начальных условиях. Для ее решения необходимо определить общую часть графов, выполнить разрез таким образом, чтобы их общая часть осталась только в одном из начальных графов, и составить таблицу соответствия (переходов) между вершинами графов для возможности осуществления переходов между ними. Изначально предполагается, что графы представлены в виде списков смежности, но в процессе работы алгоритма они преобразовываются в хеш-таблицы для быстрого определения общей части графов, которое реализуется при помощи обхода одного из графов и проверки наличия вершин во втором. Составление таблицы переходов между графами осуществляется при помощи обхода графа по парам «родительская-дочерняя» вершины, в ходе которой проверяются условия добавления узлов в таблицу переходов. Предлагается алгоритм решения задачи объединения частей ориентированного графа и приведен пример его использования.

Ключевые слова: графы, программное обеспечение, алгоритмы, объединение графов, разрез графа, разрезающее множество.

Voevoda A.A., Romannikov D.O. **Algorithm of Uniting of Parts of Oriented Graph.**

Abstract. The task of uniting graphs with a common part that were received as the result of series of simulations of a Petri net with using of program package Colored Petri Nets Tools in which a process address space is restricted by 2^{32} bytes starting with different vertices with different initial conditions is considered. For its solving it is necessary to determine the graphs common part, to perform graphs cutting in such a way that their common part will be only one of the initial graphs, and compose a table of accordance (transitions) between the graphs vertices for possibility of making the transitions between them. Firstly, we assume that the graphs are represented in form of adjacency lists, but they are converted into hash tables during the algorithm work. It's required for fast determination of the common part of the graphs that are implemented with help of traversing one of the graph and checking that the nodes exist in the second graph. Composing of the transition table is realized with help of graph traversal by "parents-child" vertex pairs and check that one of the nodes of pair can be added to the table. The algorithm for solving the problem of uniting the parts of directed graph is offered, and an example of its use.

Keywords: graphs, software, algorithms, graphs union, cutting graph, cutting set.

1. Введение. В настоящее время отсутствие критических ошибок в пользовательских сценариях программного обеспечения (ПО) на практике решается в основном за счет использования методологических способов [1], которые позволяют найти наиболее простые ошибки в основных сценариях использования ПО в лабораторной обстановке. Формальные инструменты анализа ПО, к которым можно отнести статические анализаторы [2–4], инструменты проверки моделей [5] и

др. [6], позволяют выявить достаточно широкий класс ошибок, включая те, которые трудно обнаружить в лабораторных условиях. Например, в работах [7-9] в качестве формального инструмента верификации используются сети Петри, с помощью которых строится модель ПО, а ее интерпретация позволяет определить некоторые классы ошибок. В приведенных работах для анализа и интерпретации модели сетей Петри используется программный пакет CPN Tools 4.0.1 (Colored Petri Nets Tools). Однако размеры решаемых задач анализа ПО превосходят возможности используемых инструментов: пакет моделирования сетей Петри CPN Tools имеет ограничение для хранения графа пространства состояний (ГПС) в 2^{32} байт из-за использования 32х битной архитектуры ML компилятора. При решении вышеприведенных задач, особенно в случаях многопоточности [7, 8], достаточно часто возникает проблема, что ГПС не помещается в память процесса в ~2 ГБ (2 ГБ для пространства ядра и 2 ГБ для пользовательского пространства). Для обхода этого ограничения можно выполнить серию дополнительных моделирований с места, где ГПС перестал умещаться в адресное пространство. Итогом серии моделирований является набор ГПС, где пары графов имеют общие части. После этого нужно удалить общие части из графов и составить таблицу переходов, дополняющую таблицы переходов начальных графов.

2. Постановка задачи. Из всей серии моделирований рассмотрим случай для двух ориентированных графов (далее просто графов) (G_i = {V_i, E_i}, где V – множество вершин, а E – множество переходов), т.к. решив данную задачу для случая из двух графов можно применить полученное решения для остальных графов из серии моделирования.

С точки зрения алгоритма не принципиально, из какого графа будет удалена общая часть, поэтому для определенности будем считать, что общая часть останется в графе G_2 и будет удалена из G_1.

После удаления общей части необходимо решить задачу обеспечения переходов между двумя графами при его обходе. Для этого нужно решить задачу построения таблицы переходов вида {«узел»: [«список переходов»], …}, где ключом таблицы является узел графа, а значением – список переходов.

Таким образом, в работе рассматривается задача для двух графов G_1 = {V_1, E_1}, G_2 = {V_2, E_2} с общей частью $V_3 \in V_1$, $V_3 \in V_2$, для которых требуется: 1) удалить общую часть V_3 из графа G_1; 2) построить таблицу переходов между двумя графами.

3. Решение. Рассмотрим рисунок 1, на котором изображены два графа: первый G_1, с вершиной (подразумевается вершина, с которой начиналось моделирование) N_{r1}, представлен в виде белых вершин со

сплошными линиями переходов, второй G_2, с вершиной N_{r2}, представлен в виде закрашенных вершин с пунктирными линиями переходов. Общая часть графов представлена в виде заштрихованных вершин. Размер общей части неизвестен. При этом существуют переходы из G_1 не только в вершину графа G_2, но и в его дочерние узлы. При этом могут возникать ситуации, когда: 1) в G_1 есть переходы в общую часть с G_2 (переход e_1); 2) из G_2 есть переходы в G_1 (внутри общей части графов G_1 и G_2) (переход e_2). Ситуация, где из графа G_2 есть переход в граф G_1 (не в общую часть), невозможна, т.к. при этом узел графа G_1, в который осуществляется переход, должен принадлежать графу G_2.

Предположим, что графы представлены в виде хеш-таблиц, где узлы являются ключами. Тогда определить общую часть графов G_1 и G_2 достаточно просто: потребуется лишь обойти один из графов и проверить наличие его узлов в другом. Можно использовать обход графа в ширину или глубину [10-13]. После этого необходимо удалить общую часть графа из графа G_1.

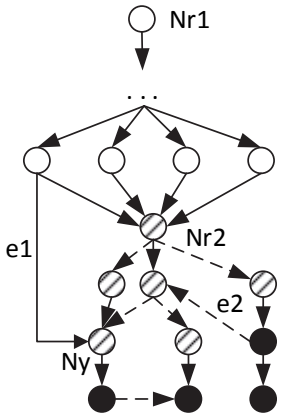


Рис. 1. Два графа с общей частью

Задача составления таблицы переходов между графами схожа с подзадачей определения минимального разреза в задаче о транспортных потоках [14, 15]. При этом таблицу переходов можно составить, найдя разрезающее множество [10] между общей частью и графом G_1. Для определения разрезающего множества необходимо выполнить обход графа G_1 и найти пары, для которых выполняются условия:

1. Существует переход в узел n графа G_2, у которого родительский узел p принадлежит графу G_1, но не принадлежит графу G_2: $e = (p, n), p \in G_1, n \in G_2, p \notin G_2$;

2. Существует переход из узла p принадлежащего общей части графов в узел n, который принадлежит только графу G_1: $e = (p, n), p \in G_1, p \in G_2, n \in G_1, n \notin G_2$.

После этого нужно добавить узел p в таблицу переходов.

4. Алгоритм. Рассмотрим реализацию алгоритма (листинг 1) для решения данной задачи и проиллюстрируем его работу на графах с рисунка 2. Раскраска для обозначения графов аналогична с графами на рисунке 1. Алгоритм приводится в виде псевдокода, в котором, будем предполагать, что переменные передаются по ссылке, т.е. их модификация внутри функции меняет значение в вызывающем коде.

Алгоритм начинает свою работу с функции main, где изначально при помощи функции makeHashMap выполняется предобработка и формируются хеш-таблицы, в которые помещаются узлы графов G_1 и G_2 в качестве ключей, путем обхода графа в ширину.

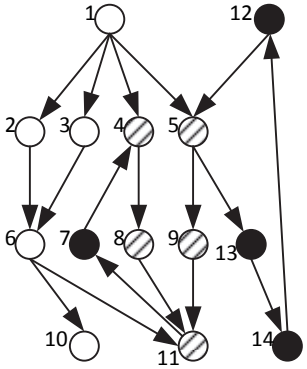


Рис. 2. Пример двух графов с общей «боковой» частью

После этого, согласно вышеприведенному описанию, определяется общая часть графов в функции findCommon (общая часть помещается в переменную graphCommon), в которой выполняется цикл по всем ключам графа G_1 из хеш-таблицы и в цикле выполняется проверка на наличие этих ключей в хеш-таблице графа G_2. Переменная graphCommon будет иметь значение {4: true, 5: true, 8: true, 9: true, 11: true}.

```
makeHashMap(graph, hashMap):
  Queue q // очередь
  q.push (вершина графа graph, с которой будем начинать обход)
  while пока в очереди есть элементы:
    Node n = q.pop()
    hashMap[n] = n
    for каждой вершины child ∈ n.children:
      if вершина child не содержится в хеш-таблице hashMap
        q.push(child)

findCommon (g1GraphHash, g2GraphHash, graphCommon):
  for каждой вершины node из ключей g1GraphHash:
    if вершина node содержится в хеш-таблице g2GraphHash:
      graphCommon[node] = true

findBoundAndRemoveCommon (g1HashMap, g2HashMap, graphCommon, bound):
  Queue q // очередь
  HashMap usedNodes // хеш-таблица для обхода графа
  q.push( пара [nil, g1HashMap.top] с которой будем начинать обход)
  while пока в очереди есть элементы:
    [Node parent, Node n] = q.pop()
    usedNodes[[parent, n]] = true
    if первое условие добавления в таблицу переходов: p ∈ G1, n ∈ G2, p ∉ G2
      bound[n] = переходы из g1HashMap[n] и g2HashMap[n]
    else if второе условие добавления в таблицу переходов: p ∈ G1, p ∈ G2, n
∈ G1, n ∉ G2
      bound[p] = переходы из g1HashMap[p] и g2HashMap[p]
    for каждой вершины child ∈ n.children:
      if пара [n, child] не содержится в хеш-таблице usedNodes:
        q.push([n, child])

  for каждой вершины node из ключей graphCommon:
    if в node из G1 есть переходы отличные от node из G2:
      g2HashMap[node] = переходы g2HashMap[node] и g1HashMap[node]
    g1HashMap.remove(node)


main():
  HashMap g1GraphHash, g2GraphHash, graphCommon, bound
  Graph g1Graph, g2Graph

  makeHashMap(g1Graph, g1GraphHash)
  makeHashMap(g2Graph, g2GraphHash)

  findCommon(g1GraphHash, g2GraphHash, graphCommon)
  findBoundAndRemoveCommon(g1GraphHash, g2GraphHash, graphCommon,
bound)
```

Листинг. 1 Алгоритм объединения частей графов

Для определения таблицы переходов для графов и удаления общей части из графа G_1 используется функция findBoundAndRemoveCommon, в которой выполняется обход графа G_1 по парам «родительский узел - дочерний узел» и каждая пара проверяется на то, является ли переход между этими узлами переходом из разрезающего множества. Дочерние узлы переходов разрезающего множества помещаются в переменную bound. В ходе обхода будут пройдены следующие пары: [nil, 1], [1, 2], [1, 3], [1, 4], [1, 5], [2, 6], [3, 6], [6, 10], [6, 11], [4, 8], [8, 11], [5, 9], [9, 11]. Среди этих пар условие добавления в хеш-таблицу срабатывает для пар [1, 4], [1, 5], [6, 11] и переменная bound будет иметь значение {4: [8], 5: [9, 13], 11: [7]}. Следует заметить, что для узлов вышеприведенной таблицы переходы состоят из переходов обоих графов.

Также в этой функции выполняется удаление общей части графов из G_1. При удалении узлов общей части из графа G_1 возможна такая ситуация, когда существуют два узла принадлежащие обоим графам (узлы 8 и 11 на рисунке 2), а переход между ними есть только в вершине графа, из которого она будет удалена (в данном случае – это граф G_1). Для предотвращения потери переходов при удалении общей части графов необходимо добавить переходы из удаляемых узлов графа G_1 в те же узлы графа G_2 (очевидно, что добавлять имеет смысл только те узлы, которых не было во втором графе).

Перейдем к анализу асимптотической сложности [10–13] алгоритма. В функции makeHashMap выполняется обход графа в ширину, сложность которого определяется как O(n), тогда для последовательных вызовов данной функции для графов G_1 и G_2 асимптотическая сложность алгоритма будет O(n + m), где n – количество узлов в графе G_1, m – в графе G_2.

В функции findCommon также выполняется обход графа G_2, что не меняет общую асимптотическую сложность.

Следующим действием является вызов функции findBoundAndRemoveCommon, в которой определяется разрезающее множество и удаляется общая часть из графа G_1. Вторая часть алгоритма также реализована на обходе графа и не меняет общую сложность. Рассмотрим часть, где выполняется определение общей границы. Данная часть алгоритма является модифицированным вариантом обхода графа в ширину за исключением того, что в обходе участвуют не узлы графа, а пары «родительский узел-дочерний узел», что меняет асимптотическую сложность алгоритма с O(n) на O (n^2). Таким образом, общая асимптотическая сложность алгоритма будет вычисляться

следующим выражением: O (n + m) + O (n) + O (n) + O (n^2), что в итоге приводит к O (n^2). Следует отметить, что вышеприведенная сложность указана для наихудшего варианта, когда все узлы графа соединены между собой. При анализе ПО таких вариантов не возникает и поэтому в решаемых задачах асимптотическая сложность алгоритма близка к O (n+m).

5. Пример. Рассмотрим граф на рисунке 3. Данный граф приводится в работах [9, 16] и использовался для анализа части программного обеспечения работы банкомата. Граф сгенерирован на основании обхода модели в сетях Петри при помощи программного пакета CPN Tools. В узлах графов (на рисунке они представлены в виде прямоугольников с закруглёнными углами) показано номер узла в верхней части, количество входных и выходных переходов в нижних левом и правом частях узла. Предположим, что он был получен по частям: сначала граф G_1, потом граф G_2. Граф G_1 задается соотношением из первой колонки таблицы 1, а граф G_2 соотношением из второй колонки. Покажем работу алгоритма на данном примере.

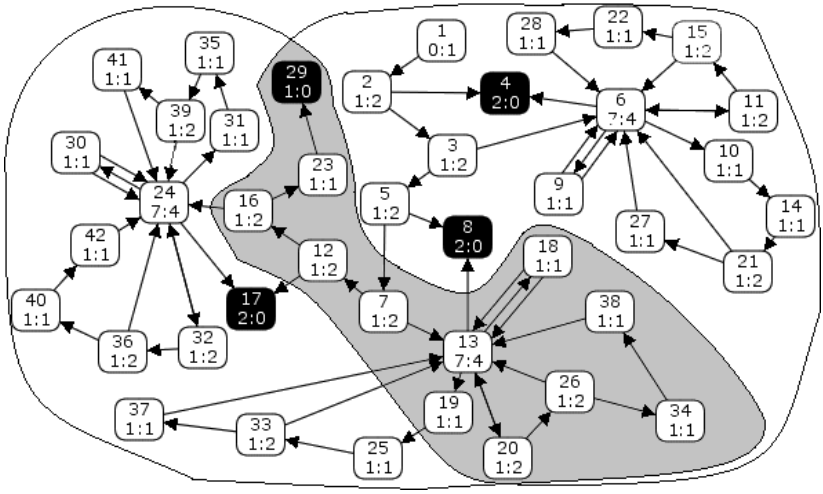


Рис. 3. Часть графа состояний сети Петри, в которой моделируется взаимодействие банкомата и пользователей

Алгоритм начинает свою работу с вызовов функций makeHashMap для обоих графов. Внутри этой функции заполняется переданная хеш-таблица: выполняется обход графа в глубину и каждый узел графа помещается в хеш-таблицу в качестве ключа и ссылка на тот же узел в качестве значения. Для рассматриваемого примера это

имеет место, и алгоритм не завершается на данном этапе. Далее следует вызов функции findCommon, где выполняется перебор всех ключей из хеш-таблицы, соответствующей первому графу, и заполняется переменной graphCommon, если узел принадлежит обоим графам. В данном примере переменная graphCommon будет иметь следующее значение: «{7: true, 12: true, 13: true, 16: true, 18: true, 19: true, 20: true, 23: true, 26: true, 29: true, 34: true, 38: true}». Последним шагом алгоритма является определение границы графов и удаление общей части.

Таблица 1. Представление графов с рисунка 3 в виде списков смежности

Граф G_1		Граф G_2	
[1: [2],	2: [3, 4],	[7: [12, 13],	12: [16, 17],
3: [5, 6],	4: [],	13: [18, 19, 20],	16: [23, 24],
5: [7, 8],	6: [9, 10, 11],	17: [],	18: [13],
7: [12, 13],	8: [],	19: [25],	20: [13, 26],
9: [6],	10: [14],	23: [29],	
11: [15],	12: [16],	24: [17, 30, 31, 32],	25: [33],
13: [8, 18, 19, 20],	14: [21],	26: [13, 34],	29: [],
15: [6, 22],	16: [23],	30: [24],	31: [35],
18: [13],	19: [],	32: [36],	33: [13, 37],
20: [13, 26],	21: [6, 27],	34: [38],	35: [39],
22: [28],	23: [29],	36: [24, 40],	37: [13],
26: [13, 34],	27: [6],	38: [13],	39: [24, 41],
28: [6],	29: [],	40: [42],	41: [24],
34: [38],	38: [13]]	42: [24]]	

При определении границы графов выполняется обход в глубину графа G_1, где для каждой пары [«родительский узел», «дочерний узел»] определяется необходимость его добавления в переменную bound, под которой понимают таблицу переходов между графами. Обход начинается с пары [nil, 1], после которой следует пары [1, 2], [2, 4] и т.д. до пары [5, 7]. Для последней пары выполняется первое условие добавления в переменную bound, которая принимает значение {7: [12, 13]}. Ключ от узла 7 в переменной bound содержит переходы, совпадающие с переходами в том же узле в графе G_1, но если бы из узла 7 был переход в граф G_2 (при этом такой переход содержался бы только в описании графа G_2), то он также бы был добавлен в список. Для пары [13, 8] также сработает условие добавления в переменную bound, которая примет окончательное значение {7: [12, 13], 13: [8, 18, 19, 20]}.

6. Заключение. В работе приведен алгоритм, позволяющий выполнить объединение двух графов с общей частью, с последующим удалением общей части из одного из графов и построение таблицы переходов для однозначных переходов между графами. Применение данного алгоритма позволяет увеличивать размер анализируемого

графа пространства состояния и, тем самым, увеличивать количество состояний в анализируемой программе.

Литература

1. *Орлов С.А.* Технология разработки программного обеспечения // Питер. 2012. 609 с.
2. *Islam S., Krinke J., Binkley D., Harman M.* Coherent clusters in source code // The Journal of Systems and Software. 2014. vol. 88. pp. 1–24.
3. *Burgstaller B., Scholz B., Blieberger J.* A symbolic analysis framework for static analysis of imperative programming languages // The Journal of Systems and Software. 2012. vol. 85. pp. 1418–1439.
4. *Аветисян А. И.* Современные методы статического и динамического анализа программ для автоматизации процессов повышения качества программного обеспечения: дисс. доктора физ. мат. наук // Москва: 2012. 271 с.
5. *Clarke E.M., O. Grumberg, D. Peled.* Model Checking // The MIT Press. 1999. 330 p.
6. *Шудрак М.О., Золотарев В.В.* Модель, алгоритмы и программный комплекс автоматизированного поиска уязвимостей в исполняемом коде // Труды СПИИРАН. 2015. Вып. 42. С. 212–231.
7. *Романников Д.О.* Разработка программного обеспечения с применением UML диаграмм и сетей Петри для систем управления локальным оборудованием дисс. канд. техн. наук // Новосибирск: 2012. 195 с.
8. *Коротиков С.В.* Применение сетей Петри в разработке программного обеспечения центров дистанционного контроля и управления: дисс. канд. техн. наук // Новосибирск: 2007. 216 с.
9. *Марков А.В.* Автоматизация проектированияи анализа программного обеспечения с использованием языка UML и сетей Петри: дисс. канд. техн. наук // Новосибирск: 2015. 176 с.
10. *Cormen T., Leiserson C., Rivest R., Stein C.* Introduction to Algorithms: 3rd Edition // The MIT Press. 2009. 1328 p.
11. *Even S.* Graph Algorithms: 2nd Edition // Cambridge University Press. 2011. 187 p.
12. *Tarjan R.* Data Structures and Network Algorithms // Society for Industrial and Applied Mathematics. 1983.
13. *Sedgewick R.* Algorithms in C++: 3rd Edition // Addison-Wesley Professional. 1998. 752 p.
14. *Goldberg A., Tardos E., Tarjan R.* Network flow algorithms // Springer. 1990. pp. 101–164.
15. *Schhrijver A.* Paths and flows – A historical survey // CWI Quarterly. 1993. pp. 169–183.
16. *Марков, А.В., Воевода А.А.* Анализ сетей Петри при помощи деревьев достижимости // Сб. науч. тр. НГТУ. 2013. №. 71. С. 78–95.

References

1. Orlov S.A. *Tehnologija razrabotki programmnogo obespechenija* [Technology of software development]. Piter. 2012. 609 p. (In Russ.).
2. Islam S., Krinke J., Binkley D. Coherent clusters in source code. *The Journal of Systems and Software*. 2014. vol. 88. pp. 1–24.
3. Burgstaller B., Scholz B., Blieberger J. A symbolic analysis framework for static analysis of imperative programming languages. *The Journal of Systems and Software*. 2012. vol. 85. pp. 1418–1439.
4. Avetisjan A. I. *Sovremennye metody staticheskogo i dinamicheskogo analiza programm dlja avtomatizacii processov povyshenija kachestva programmnogo obespechenija: diss. doktora. fiz. mat. Nauk* [Modern methods of static and dynamic

analysis software for process automation to improve the quality of software: doctor phys. math. thesis]. Moskow: 2012. 271 p. (In Russ.).
5. Clarke E.M., O. Grumberg, D. Peled. Model Checking. The MIT Press. 1999. 330 p.
6. Shudrak M.O., Zolotarev V.V. [Models, algorithms and software for automated vulnerability scan executable code]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2015. vol. 42. pp. 212–231.
7. Romannikov D.O. *Razrabotka programmnogo obespechenija s primeneniem UML diagramm i setej Petri dlja sistem upravlenija lokal'nym oborudovaniem: diss. kand. tehn. nauk* [Software development using UML diagrams and Petri nets for local control systems equipment Ph.D. thesis]. Novosibirsk: 2012. 195 p. (In Russ.).
8. Korotikov S.V. *Primenenie setej Petri v razrabotke programmnogo obespechenija centrov distancionnogo kontrolja i upravlenija diss. kand. tehn. nauk* [Application of Petri nets in software development centers, remote monitoring and control: Ph.D. thesis]. Novosibirsk: 2015. 216 p. (In Russ.).
9. Markov A. V. *Avtomatizacija proektirovanijai analiza programmnogo obespechenijas ispol'zovaniem jazyka UML i setej Petri: diss. kand. tehn. nauk* [Computer-aided design and analysis software with UML and Petri nets: Ph.D. thesis]. Novosibirsk: 2007. 176 p. (In Russ.).
10. Cormen T., Leiserson C., Rivest R., Stein C. Introduction to Algorithms: 3rd Edition. The MIT Press. 2009. 1328 p.
11. Even S. Graph Algorithms: 2nd Edition. Cambridge University Press. 2011. 187 p.
12. Tarjan R. Data Structures and Network Algorithms. Society for Industrial and Applied Mathematics, 1983.
13. Sedgewick R. Algorithms in C++: 3rd Edition. Addison-Wesley Professional. 1998. 752 p.
14. Goldberg A., Tardos E., Tarjan R. Network flow algorithms. Springer. 1990. pp. 101–164.
15. Schhrijver A. Paths and flows – A historical survey. CWI Quarterly. 1993. pp. 169–183.
16. Markov A.B., Voevoda A.A. [Petri Nets Analysis with help of reachability trees]. *Sbornik nauchnykh trudov NGTU – Collection of scientific papers of NSTU*. 2013. vol. 1 (71). pp. 78–95. (In Russ.).

Романников Дмитрий Олегович — к-т техн. наук, доцент, доцент кафедры автоматики, Новосибирский государственный технический университет. Область научных интересов: верификация, анализ программ. Число научных публикаций — 40. rom2006@gmail.com; пр. Карла Маркса 20, Новосибирск, 630073; р.т.: +7 961 223 8567.

Romannikov Dmitry Olegovich — Ph.D., associate professor, associate professor of automation department, Novosibirsk State Technical University. Research interests: verification, program analysis. The number of publications — 40. rom2006@gmail.com; 20, Karl Marx Avenue, Novosibirsk, 630073; office phone: +7 961 223 8567.

Воевода Александр Александрович — д-р техн. наук, профессор, профессор кафедры автоматики, Новосибирский государственный технический университет. Область научных интересов: полиномиальный синтез, сети Петри, UML диаграммы. Число научных публикаций — 200. voevoda@ucit.ru; пр. Карла Маркса 20, Новосибирск, 630073; р.т.: +79139223092.

Voevoda Alexandr Aleksandrovich — Ph.D., Dr. Sci., professor, professor of automation department, Novosibirsk State Technical University. Research interests: polynomial synthesis, UML diagrams, Petri nets. The number of publications — 200. voevoda@ucit.ru; 20, Karl Marx Avenue, Novosibirsk, 630073; office phone: +79139223092.

РЕФЕРАТ

Воевода А.А., Романников Д.О. **Алгоритм объединения частей ориентированного графа.**

Рассматривается задача объединения частей графа, которые содержат общую частью. Данные части графов могут быть получены различными способами, например, при моделировании сети Петри с использованием программного пакета CPN Tools (в котором есть ограничение адресное пространство процесса в 2^{32} байт из-за использования 32-х битного компилятора языка ML). Для решения задачи объединения частей графа необходимо выполнить следующее: 1) удалить общую часть графов; 2) составить таблицу переходов между графами состоящую из вершин графов для возможности выполнять переходы между частями графов.

В работе предложен алгоритм, в котором предполагается, что изначально графы, представленные в виде списков смежности, преобразуются в хеш-таблицы. Определение общей части графов выполняется с помощью обхода одного из графов и проверки на вхождение его узлов в другой. Составление таблицы переходов между графами осуществляется при помощи обхода графа по парам «родительский - дочерний» узел, в ходе которой проверяется условия добавления узлов в таблицу переходов.

В работе так же оценена асимптотическая сложность алгоритма и представлены примеры его применения. Полученное решение в дальнейшем может быть последовательно применено для объединения всего множества графов.

SUMMARY

Voevoda A.A., Romannikov D.O. **Algorithm of Uniting of Parts of Oriented Graph.**

The problem of uniting parts of the graph, which contains a general part. The given parts of the graphs can be produced by various methods, for example, in a Petri net modeling with using of a software package CPN Tools (which has a limited address space of 2^{32} bytes in cause of using of 32-bit compiler of ML language). To solve the task of uniting parts of the graph it's need to do the following: 1) remove the general part of the graphs; 2) create a table of transitions between graphs consisting of vertices of graphs to be able to perform transitions between parts of the graph.

The algorithm, which assumes that initially presented as graphs adjacency lists are converted into hash tables. Determination of the total of the graph traversal is performed via one of the graphs and check his entry node to another. Making the transition table between the graphs made using graph traversal pairs "parent – child" node in order to verify the conditions of adding nodes to a jump table.

The paper also evaluated the asymptotic complexity of the algorithm and provides examples of its use. The resulting solution can then be applied consistently to unite the whole set of the graphs.

УДК 004.93.51

Т.М. КОСОВСКАЯ

САМООБУЧАЮЩАЯСЯ СЕТЬ С ЯЧЕЙКАМИ, РЕАЛИЗУЮЩИМИ ПРЕДИКАТНЫЕ ФОРМУЛЫ

Косовская Т.М. **Самообучающаяся сеть с ячейками, реализующими предикатные формулы.**

Аннотация. Рассматривается модель перенастраиваемой сети с ячейками, реализующими предикатные формулы, имеющие вид элементарных конъюнкций. В отличие от классических нейронных сетей предлагаемая модель имеет два блока: блок обучения и блок решения. При ошибках, возникающих при использовании блока решения, подключается блок обучения. Кроме того, конфигурация сети не фиксируется заранее, а меняется каждый раз после работы блока обучения. Базой для создания перенастраиваемой логико-предикатной сети является логико-предметный подход к решению задач искусственного интеллекта, а также понятие неполной выводимости предикатной формулы, позволяющее выделять общие подформулы элементарных конъюнкций.

Ключевые слова: искусственный интеллект, формула исчисления предикатов, уровневое описание классов, самообучающаяся распознающая сеть.

Kosovskaya T.M. **Self-training Network with the Sells Implementing Predicate Formulas.**

Abstract. A model of self-modificated predicate network with cells implementing predicate formulas in the form of elementary conjunction is suggested. Unlike a classical neuron network the proposed model has two blocks: a training block and a recognition block. If a recognition block has a mistake then the control is transfered to a training block. Always after a training block implementation the configuration of a recognition block is changed. The base of the proposed logic-predicate network is a logic-objective approach to AI problems solving and level description of classes as well as the notion of partial deducibility which allows to extract common sub-formulas of elementary conjunctions.

Keywords: artificial intelligence, pattern recognition, predicate calculus formulas, level description of a class, self-training recognition network

1. Введение. Традиционно при моделировании задач искусственного интеллекта (ИИ), а особенно задач распознавания образов, исследуемый объект рассматривается как неделимое целое и описывается глобальными признаками, характеризующими его свойства. Такой подход плохо приспособлен к моделированию сложных объектов, характеризующихся свойствами его элементов и отношениями между ними.

40 лет назад появилось большое количество монографий с одним и тем же названием «Искусственный интеллект» (среди них, например, книга с другим названием [1]), в которых предлагалось использование языка исчисления предикатов (ИП) и автоматического доказательства теорем с помощью метода резолюций для решения разнообразных за-

дач этой тематики. Язык ИП [2] вполне адекватен для моделирования сложных и изменяющихся объектов. Однако в этих монографиях не были сделаны оценки числа шагов решения задач в такой модели, что не позволило применять её на практике. В 2006 году вышел перевод монографии почти с таким же названием [10], в которой вновь предлагается использование языка ИП. Из оценок числа шагов приводится лишь экспоненциальная зависимость длины описания объекта в виде строки некоторых значений от его описания на языке ИП.

В работе автора [4] доказаны оценки числа шагов алгоритмов, решающих задачи ИИ при их моделировании с помощью языка ИП. Анализ этих оценок позволил разработать иерархические многоуровневые описания [6] целевых условий, существенно уменьшающие время решения задач. На основе многоуровневых описаний в [5] было предложено построение нейронной сети. Однако на тот момент методика обучения такой сети не была разработана. Возможность автоматического создания логико-предметной сети по обучающей выборке появилась после разработки алгоритма построения многоуровневого описания классов [9].

Широко распространена модель, в которой элемент классической искусственной нейронной сети [3] представляет из себя сумматор взвешенных входов, после которого находится передаточная функция, приводящая значение выхода сумматора в промежуток $[0, 1]$. Конфигурация нейронной сети заранее фиксируется и в процессе обучения меняются только значения весов входов сумматора.

Ниже предлагается модель логико-предикатной нейронной сети, имеющей два блока: блок обучения и блок распознавания. Каждый из блоков в качестве своих элементов имеет предикатную формулу в виде элементарной конъюнкции. Входами элемента сети являются значения предметных переменных для соответствующей элементарной конъюнкции и значения атомарных предикатных формул, задающих свойства предметных переменных и отношения между ними.

Конфигурация блока обучения формируется в процессе обучения сети. После предварительного обучения в этом блоке определяется конфигурация блока распознавания. Блок обучения — это «долго работающий» блок. В отличие от него блок распознавания — это «быстро работающий» блок. Несмотря на то, что блок обучения работает действительно долго (решается NP-трудная задача), это соответствует тому, что человек обучается годами, чтобы потом решать многие задачи в течение секунд.

2. Общая постановка задачи. Пусть исследуемый объект представлен как множество своих элементов $\omega = \{\omega_1, \ldots, \omega_t\}$. На ω задан набор предикатов $p_1, \ldots, p_n$, характеризующих свойства элементов ω и отношения между ними. Логическим описанием $S(\omega)$ объекта ω называется множество всех атомарных формул или их отрицаний, истинных на ω. Множество всех объектов разбито на классы $\Omega = \cup_{k=1}^{K}\Omega_k$. Логическим описанием класса Ω_k называется формула $A_k(\overline{x})$, заданная в виде дизъюнкции элементарных конъюнкций, такая что если $A_k(\overline{\omega})$ истинна, то $\omega \in \Omega_k$.[1]

С помощью построенных описаний объектов и классов в [8] предлагается решать следующие задачи.

Задача идентификации. *Проверить, удовлетворяет ли объект ω или его часть описанию класса $A_k(\overline{x})$ и предъявить эту часть объекта.*

Задача классификации. *Найти все такие номера k, что верна формула $A_k(\overline{\omega})$.*

Задача анализа. *Найти и классифицировать все части τ объекта ω, для которых $A_k(\overline{\tau})$.*

Решение задач идентификации, классификации и анализа для распознавания сложного объекта сведено в к доказательству соответственно логических следований[2]

$$S(\omega) \Rightarrow \exists \overline{x}_{\neq}\ A_k(\overline{x}), \tag{1}$$

$$S(\omega) \Rightarrow \bigvee_{k=1}^{M} A_k(\overline{\omega}), \tag{2}$$

$$S(\omega) \Rightarrow \bigvee_{k=1}^{M} \exists \overline{x}_{\neq}\ A_k(\overline{x}). \tag{3}$$

[1] Здесь и далее посредством $\overline{x}$ обозначается список элементов конечного множества x, соответствующий некоторой перестановке номеров его элементов. Тот факт, что элементами списка $\overline{x}$ являются элементы множества y, будем записывать в виде $x \subseteq y$.

[2] Для того, чтобы записать, что значения для переменных списка $\overline{x}$, удовлетворяющие формуле $A(\overline{x})$, различны, вместо формулы

$$\exists x_1 ... \exists x_m (\&_{i=1}^{m} \&_{j=i+1}^{m} (x_i \neq x_j)\ \&\ A(x_1, ..., x_m))$$

будет использоваться обозначение

$$\exists \overline{x}_{\neq} A(\overline{x}).$$

Строго говоря, вместо формул (1), (2), (3) следовало бы писать соответственно

$$S(\omega) \Rightarrow (?\overline{x}_{\neq})\ A_k(\overline{x}), \qquad (1')$$

$$S(\omega) \Rightarrow (?k_{k=1}^{M})A_k(\overline{\omega}), \qquad (2')$$

$$S(\omega) \Rightarrow (?k_{k=1}^{M})(?\overline{x}_{\neq})\ A_k(\overline{x}), \qquad (3')$$

но рассматриваемые алгоритмы доказательства логических следований не только отвечают на вопрос *«существвует ли ... ?»*, но и предъявляют значения для переменных [2].

Заметим, что для того, чтобы уметь доказывать (1), (2), (3), достаточно уметь доказывать логическое следование

$$S(\omega) \Rightarrow \exists\overline{x}_{\neq}\ A(\overline{x}), \qquad (4)$$

где $A(\overline{x})$ — элементарная конъюнкция атомарных формул и их отрицаний. В [4,8] доказаны оценки числа шагов алгоритмов, решающих задачу (4), а также задачи (1), (2), (3). Эти оценки имеют экспоненциальный от длины записи формулы $A(\overline{x})$ вид. Для алгоритма полного перебора в показателе оценки находится количество переменных формулы $A(\overline{x})$, а для алгоритмов, основанных на построении вывода в исчислении предикатов, в показателе оценки находится количество атомарных формул, входящих в формулу $A(\overline{x})$.

Доказана NP-полнота задач (1), (2), (3) и, следовательно, NP-трудность задач (1'), (2'), (3').

3. Многоуровневое описание классов. Для уменьшения числа шагов работы алгоритмов, решающих описанные задачи, в [6] предложено многоуровневое описание классов распознаваемых объектов, по сути своей являющееся иерархическим описанием классов и учитывающее составляющие конструкции объектов. В [5] описана возможность построения логико-предикатной нейронной сети на основании уже имеющегося многуровневого описания классов.

Алгоритм автоматического построения многуровневого описания класса, позволяющий выделить обобщённые характеристики объектов, присущие объектам одного класса, описан в [9]. Этот алгоритм базируется на понятии неполной выводимости предикатной формулы, описанном в [7].

Рассматриваются объекты, структура которых позволяет выделить достаточно простые их части и дать описание объекта в терминах свойств этих частей и отношений между ними. В частности, это можно

сделать, выделяя «часто» встречающиеся подформулы $P_i^1(\overline{y}_i^1)$, формул $A_k(\overline{x})$ «небольшой сложности». При этом записывается система равносильностей вида $p_i{}^1(y^1) \Leftrightarrow P_i^1(\overline{y}_i^1)$, где $p_i{}^1$ – новые предикаты, которые будем называть предикатами 1-го уровня, а переменные y_i^1 – новые переменные для списков исходных переменных, которые будем называть переменными 1-го уровня.

Обозначим формулы, полученные из $A_k(\overline{x}_k)$ путем замены всех вхождений формул вида $P_i^1(\overline{y}_i^1)$ на атомарные формулы $p_i^1(x_i^1)$ (при $y_i^1 \subseteq x$) посредством $A_k^1(\overline{x}_k^1)$. Здесь $\overline{x}_k^1$ – список всех переменных формулы $A_k^1(\overline{x}^1)$, состоящий как из некоторых (быть может всех) исходных переменных формулы $A_k(\overline{x}_k)$, так и из переменных первого уровня, появившихся в формуле $A_k^1(\overline{x}_k^1)$. Такие формулы $A_k^1(\overline{x}_k^1)$ можно рассматривать как описания классов в терминах предикатов исходного (нулевого) и первого уровней.

Описанием объекта $S^1(\omega)$ первого уровня назовем множество всех атомарных формул вида $p_i^1(\omega_{ij}^1)$, для которых истинна определяющая подформула $P_i^1(\overline{\tau}_{ij}^1)$ при $\tau_{ij}^1 \subset \omega$, а объект первого уровня ω_{ij}^1 представляет из себя список исходных объектов $\overline{\tau}_{ij}^1$.

Процедуру выделения «часто» встречающихся подформул «небольшой сложности» можно повторить с формулами $A_k^1(\overline{x}^1)$.

В результате построения составных предикатов (т.е. предикатов различных уровней) и многоуровневого описания классов исходное множество описаний классов $\{A_k(\overline{x})\}$ может быть записано с помощью равносильной ей многоуровневой системы описаний классов вида

$$\left\{\begin{array}{ccc} & A_k^L(\overline{x}^L) & \\ p_1^1(x_1^1) & \Leftrightarrow & P_1^1(\overline{y}_1^1) \\ & \vdots & \\ p_{n_1}^1(x_{n_1}^1) & \Leftrightarrow & P_{n_1}^1(\overline{y}_{n_1}^1) \\ & \vdots & \\ p_i^l(x_i^l) & \Leftrightarrow & P_i^l(\overline{y}_i^l) \\ & \vdots & \\ p_{n_L}^L(x_{n_L}^L) & \Leftrightarrow & P_{n_L}^L(\overline{y}_{n_L}^L) \end{array}\right. .$$

Алгоритм многоуровневого распознавания.

Проверка следования (4) при использовании L-уровневого описания разбивается на последовательное в цикле при $l = 1, ..., L$ выпол-

нение п.п. 1 – 4 с последующим выполнением п. 5.

1. Проверка следований $S^{l-1}(\omega) \Rightarrow \exists \overline{y}^l_{i \neq} P^l_i(\overline{y}^l_i)$ $(i = 1, ..., n^l_i)$ с нахождением тех наборов $\overline{\tau}^l_i$ значений исходных констант для списка переменных $\overline{y}^l_i$, при которых $S^{l-1}(\omega) \Rightarrow P^l_i(\overline{\tau}^l_i)$ $(i = 1, ..., n^l_i)$.

2. Введение новых атомарных формул $p^l_i(y^l_i)$ l-го уровня, определяемых равносильностями $p^l_i(y^l_i) \Leftrightarrow P^l_i(\overline{y}^l_i)$ с новыми переменными y^l_i l-го уровня для списков переменных $\overline{y}^l_i$.

3. Замена в формуле $A^{l-1}(\overline{x}^{l-1})$ всех подформул вида $P^l_i(\overline{y}^l_i)$ $(i = 1, ..., n^l_i)$ на атомарные формулы $p^l_i(y^l_i)$ и получение формулы $A^l(\overline{x}^l)$.

4. Добавление в $S^{l-1}(\omega)$ постоянных атомарных формул l-го уровня вида $p^l_i(\tau^l_i)$, где τ^l_i — новые константы, задающие списки констант $\overline{\tau}^l_i$ $(i = 1, ..., n^l_i)$ и получение описания объекта l-го уровня $S^l(\omega)$.

5. Проверка следования $S^L(\omega) \Rightarrow \exists \overline{x}^L_{\neq} A^L(\overline{x}^L)$ с нахождением тех наборов $\overline{\tau}^L$ значений исходных констант для списка переменных $\overline{x}^L$, при которых $S^L(\omega) \Rightarrow A^L(\overline{\tau}^L)$.

В [8] доказаны оценки изменения числа шагов проверки (4) при использовании двухуровневого описания классов и приведены модельные примеры, иллюстрирующие существенное уменьшение показателя экспоненты при его использовании. Однако там применяется эвристическое выделение общих подформул для построения двухуровневого описания классов. В [9] описан алгоритм построения многоуровневого описания класса по обучающей выборке, в основе которого лежит понятие неполной выводимости.

4. Понятие неполной выводимости формулы. Понятие неполной выводимости предикатной формулы было введено в [7] для распознавания объектов с неполной информацией.

Рассматривается задача проверки того, что из истинности всех формул множества $S(\omega)$ следует истинность $A(\overline{x})$ или некоторой её максимальной подформулы $\widetilde{A}(\overline{y})$ на наборе различных констант из ω, где список переменных $\overline{y}$ является подсписком списка переменных $\overline{x}$.

Пусть a и $\widetilde{a}$ – количества атомарных формул в элементарных конъюнкциях $A(\overline{x})$ и в $\widetilde{A}(\overline{y})$ соответственно, m и $\widetilde{m}$ – количества предметных переменных в $A(\overline{x})$ и $\widetilde{A}(\overline{y})$ соответственно.

Числа q и r вычисляются по формулам $q = \frac{\widetilde{a}}{a}$, $r = \frac{\widetilde{m}}{m}$ и характеризуют степень совпадения формул $A(\overline{x})$ и $\widetilde{A}(\overline{y})$. В этом случае подформула $\widetilde{A}(\overline{y})$ называется (q, r)-фрагментом формулы $A(\overline{x})$.

Подформула $\widetilde{A}(\overline{y})$ называется максимальной подформулой эле-

ментарной конъюнкции $A(\overline{x})$, если она является её (q,r)-фрагментом с максимальным среди всех (q,r)-фрагментов значением параметра q. То есть для $\widetilde{A}(\overline{y})$ справедливо $S(\omega) \Rightarrow \exists\overline{y}_{\neq}\widetilde{A}(\overline{y})$ и ни для какой подформулы формулы $A(\overline{x})$, с большим значением параметра q, это следствие не выполняется.

Возможно другое определение значений параметров q и r. Пусть предикатным символам, задающим признаки объектов, приписаны веса w_i $(i = 1,\ldots,n)$, а предметным переменным, входящим в формулы, приписаны соответственно веса v_j $(j = 1,\ldots,m)$ и $\widetilde{v}_j$ $(j = 1,\ldots,m')$. Тогда $q_w = \frac{\widetilde{w}}{w}$, $r_v = \frac{\widetilde{v}}{v}$, где w и $\widetilde{w}$ – суммы весов предикатных формул в $A(\overline{x})$ и в $\widetilde{A}(\overline{y})$ соответственно, v и $\widetilde{v}$ – суммы весов предметных переменных в $A(\overline{x})$ и $\widetilde{A}(\overline{y})$ соответственно.

Параметр q, так же как и параметр q_w, характеризует, насколько информативен фрагмент, содержащий лишь r-ую (r_v-ую) часть переменных.

Задача нахождения максимального (q,r)-фрагмента формулы $\widetilde{A}(\overline{y})$ при условии справедливости множества постоянных атомарных формул $S(\omega)$ называется задачей проверки неполной выводимости этой формулы. В [7] приведён один из возможных алгоритмов её решения.

5. Нахождение наибольшей общей подформулы двух формул. Понятие неполной выводимости из множества постоянных атомарных формул или их отрицаний легко обобщается до понятия неполной выводимости двух элементарных конъюнкций.

Пусть $A(\overline{x})$ и $B(\overline{y})$ – две элементарные конъюнкции предикатных формул со списками предметных переменных $\overline{x}$ и $\overline{y}$ соответственно. Проверка неполной выводимости $A(\overline{x}) \Rightarrow_P \exists\overline{y}_{\neq}B(\overline{y})$ заключается в нахождении такого максимального (q,r)-фрагмента $Q_{AB}(\overline{z})$ формулы $B(\overline{y})$ и такой подстановки $\lambda_{AQ} = |^{z}_{y'}$ списка переменных $\overline{y'}$ из $\overline{y}$ вместо переменных списка $\overline{z}$, что $Q_{AB}(\overline{y'})$ является максимальной подформулой формулы $B(\overline{y})$, такой что $A(\overline{x}) \Rightarrow \exists\overline{y'}_{\neq}Q_{AB}(\overline{y'})$.

Аналогично при проверке неполной выводимости $B(\overline{y}) \Rightarrow_P \exists\overline{x}_{\neq}A(\overline{x})$ получаем максимальный (q,r)-фрагмент $Q_{BA}(\overline{z})$ формулы $A(\overline{x})$ и такую подстановку $\lambda_{BQ} = |^{z}_{x'}$ списка переменных $\overline{x'}$ из $\overline{x}$ вместо переменных списка $\overline{z}$, что $Q_{BA}(\overline{y'})$ является максимальной подформулой формулы $A(\overline{x})$, такой что $B(\overline{y}) \Rightarrow \exists\overline{x'}_{\neq}Q_{BA}(\overline{x'})$.

Можно доказать, что формулы $Q_{AB}(\overline{y'})$ и $Q_{BA}(\overline{x'})$ совпадают с точностью до имён переменных. В качестве максимальной (с точностью до имён переменных) общей подформулы двух элементарных

конъюнкций $A(\overline{x})$ и $B(\overline{y})$ можно взять любую из них. Обозначим такую подформулу посредством $Q(\overline{z})$.

Найденные в процессе проверки неполной выводимости подстановки λ_{AQ} и λ_{BQ} обеспечивают возможность такого переименования переменных из $\overline{z}$, что $Q(\overline{z})$ становится в точности общей подформулой элементарных конъюнкций $A(\overline{x})$ и $B(\overline{y})$ соответственно. Эти подстановки назовём унификаторами формулы $Q(\overline{z})$ с элементарными конъюнкциями $A(\overline{x})$ и $B(\overline{y})$.

6. Построение многоуровневого описания классов. Понятие неполной выводимости формулы позволяет разработать подход к выделению подформул с требуемыми свойствами [9].

Алгоритм построения многоуровневого описания.

1. Для каждой пары элементарных конъюнкций $A_i(\overline{x}_i)$ и $A_j(\overline{x}_j)$, входящих в описания классов, посредством проверки неполной выводимости для $A_i(\overline{x}_i) \Rightarrow_P \exists \overline{x}_{j \neq} A_j(\overline{x}_j)$ выделяем их максимальную (с точностью до имён предметных переменных) подформулу $Q^1_{ij}(\overline{x}_{ij})$.

При использовании как алгоритма полного перебора, так и алгоритма, основанного на построении вывода в исчислении предикатов, найденная формула $Q^1_{ij}(\overline{x}_{ij})$ является в точности подформулой элементарной конъюнкции $A_j(\overline{x}_j)$, поэтому унификатор λ_{iQ} является тождественной подстановкой. При этом будет найден унификатор λ_{jQ}.

2. Повторяем процесс выделения общих подформул для $Q^{l-1}_{i_1 \ldots i_{2^{l-1}}}(\overline{x}_{i_1 \ldots i_{2^{l-1}}})$ и $Q^{l-1}_{j_1 \ldots j_{2^{l-1}}}(\overline{x}_{j_1 \ldots j_{2^{l-1}}})$, получив их общие (с точностью до имён предметных переменных) подформулы $Q^l_{i_1 \ldots i_{2^{l-1}} j_1 \ldots j_{2^{l-1}}}(\overline{x}_{i_1 \ldots i_{2^{l-1}} j_1 \ldots j_{2^{l-1}}})$ $(l = 2, \ldots, L)$ и унификаторы для соответствующих подформул. Процесс завершится, так как на каждой итерации длины подформул уменьшаются.

3. Выберем среди подформул $Q^l_{i_1 \ldots i_{2^{l-1}} j_1 \ldots j_{2^{l-1}}}(\overline{x}_{i_1 \ldots i_{2^{l-1}} j_1 \ldots j_{2^{l-1}}})$ минимальные (по числу переменных для применения алгоритма полного перебора или по числу атомарных формул для применения алгоритмов, основанных на построении вывода в исчислении предикатов) и обозначим их посредством $P^1_i(\overline{y}^1_i)$ $(i = 1, \ldots, n_1)$.

4. Формулы $P^{l+1}_i(\overline{y}^{l+1}_i)$ $(i = 1, \ldots, n_{l+1}, l = 2, \ldots, L)$ строятся из выделенных ранее подформул $Q^l_{i_1 \ldots i_{2^{l-1}} j_1 \ldots j_{2^{l-1}}}(\overline{x}_{i_1 \ldots i_{2^{l-1}} j_1 \ldots j_{2^{l-1}}})$ с учётом минимизации требуемых параметров и того, что подформулы вида $P^1_i(\overline{y}^1_i)$ в них заменены на новые атомарные формулы $p^1_i(y^1_i)$, определяемые равносильностями $p^1_i(y^1_i) \Leftrightarrow P^1_i(\overline{y}^1_i)$.

7. Формирование логико-предикатной сети. На стадии обучения для формирования обучающего блока сети предлагается обучающая выборка, содержащая описания объектов с указанием классов, которым они принадлежат. В описании каждого объекта различные константы заменяются различными переменными и между атомарными формулами ставится знак &.

Описанием класса служит дизъюнкция так полученных элементарных конъюнкций. Затем используется *Алгоритм построения многоуровневого описания* из конъюнкций, соответствующих объектам обучающей выборки.

Полученные в обучающем блоке формулы $P_i^1(\overline{y}_i^1)$ ($i = 1, \dots n_l$, $l = 1, \dots, L$) служат содержимым ячеек l-го уровня решающего блока. Последний уровень составляют формулы $A_k^L(\overline{x_k^L})$.

В процессе распознавания используется только решающий блок. Он работает в соответствии с *Алгоритмом многоуровневого распознавания*. Если в процессе использования построенной сети обнаруживается неправильное распознавание объекта, то возможно дообучение сети посредством добавления описания неправильно классифицированного объекта к первому слою обучающего блока и выделению общих подформул этого описания из уже имеющихся. После этого происходит перестройка решающего блока. Схема логико-предикатной сети представлена на рисуноке 1.

8. Модельный пример формирования сети. Пример контурных изображений взят из [1].

Пусть дана обучающая выборка класса контурных изображений «ящиков», представленная на рисуноке 2.

Каждый объект – это совокупность вершин в контурном изображении. Описания объектов и классов объектов заданы в терминах исходных предикатов V и L, определённых на рисуноке 3 и задающих отношения между вершинами.

Описание класса «ящиков» содержит 4 элементарные конъюнкции, каждая из которых состоит из всех атомарных формул, истинных для соответствующего изображения, в которых имя вершины i заменено на переменную x_i. Так, например, изображение b на рисуноке 2 описывает элементарная конъюнкция $A_2(x_1, \dots, x_8) =$

$$V(x_1, x_4, x_2)\&V(x_2, x_1, x_6)\&V(x_2, x_6, x_3)\&V(x_2, x_1, x_3)\&$$

$$V(x_3, x_2, x_8)\&V(x_4, x_5, x_1)\&V(x_4, x_6, x_1)\&V(x_4, x_7, x_5)\&$$

$$V(x_4,x_7,x_6)\&V(x_4,x_7,x_1)\&V(x_5,x_4,x_7)\&V(x_5,x_7,x_6)\&$$
$$V(x_6,x_2,x_5)\&V(x_6,x_2,x_4)\&V(x_6,x_5,x_8)\&V(x_6,x_4,x_8)\&$$
$$V(x_6,x_8,x_2)\&V(x_7,x_5,x_4)\&V(x_7,x_8,x_5)\&V(x_7,x_8,x_4\&$$
$$V(x_8,x_3,x_6)\&V(x_8,x_6,x_7)\&V(x_8,x_3,x_7)\&L(x_5,x_4,x_6).$$

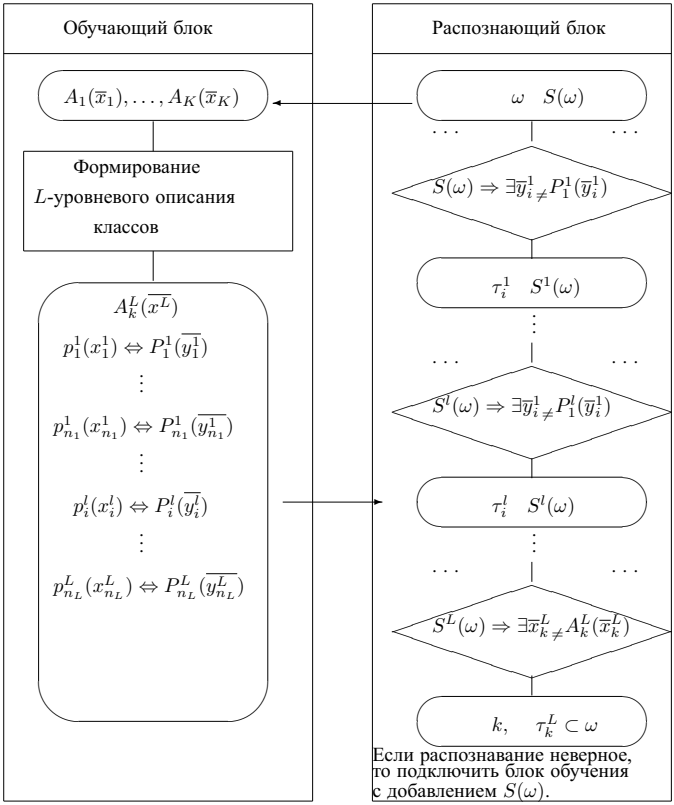


Рис. 1. Схема логико-предикатной сети

Попарная проверка частичной выводимости между этими элементарными конъюнкциями позволяет выделить их максимальные общие подформулы, соответствующие изображениям на рисунке 4.

В процессе попарного выделения общих подформул формул $A_i(\overline{x}_i)$ и $A_j(\overline{x}_j)$ $(i=1,...,3,\ j=i+1,\ldots,4)$ получили их максимальные (с точностью до имён переменных) подформулы, соответствующие

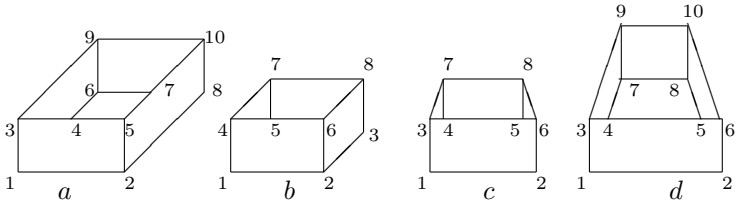
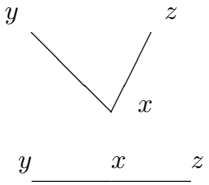


Рис. 2. Обучающая выборка



$V(x,y,z) \Longleftrightarrow (\angle yxz < \pi)$

$L(x,y,z,) \Longleftrightarrow x$ лежит на отрезке с концами y и z

Рис. 3. Исходные предикаты

изображениям, представленным на рисунке 4. Так, например, подформула $Q^1_{2,4}(x_2,x_3,x_4,x_5,x_9,x_{10}) =$

$$V(x_1,x_3,x_2)\&V(x_2,x_1,x_5)\&V(x_3,x_4,x_1)\&V(x_3,x_5,x_1)\&$$

$$V(x_3,x_9,x_4)\&V(x_3,x_9,x_5)\&V(x_3,x_9,x_1)\&V(x_5,x_2,x_4)\&$$

$$V(x_5,x_2,x_3)\&V(x_9,x_{10},x_4)\&V(x_9,x_4,x_3)\&L(x_4,x_3,x_5)$$

описывает изображение bd на рисунке 4.

После попарного выделения максимальных общих подформул формул $Q^1_{i_1,i_2}(x^1_{i_1,i_2})$ и $Q^1_{j_1,j_2}(x^1_{j_1,j_2})$ получена единственная их максимальная общая подформула $Q^2(x_1,x_2,x_3,x_4,x_5,x_9,x_{10}) =$

$$V(x_1,x_3,x_2)\&V(x_2,x_1,x_5)\&V(x_3,x_4,x_1)\&V(x_3,x_5,x_1)\&$$

$$V(x_3,x_9,x_4)\&V(x_3,x_9,x_5)\&V(x_3,x_9,x_1)\&V(x_5,x_2,x_4)\&$$

$$V(x_5,x_2,x_3)\&V(x_9,x_{10},x_3)\&L(x_4,x_3,x_5),$$

соответствующая изображению ad на рисунке 4.

Элементарная конъюнкция $P^1(x_1,x_2,x_3,x_4,x_5,x_9,x_{10}) = V(x_1,x_3,x_2)\&V(x_2,x_1,x_5)\&V(x_3,x_4,x_1)\&V(x_3,x_5,x_1)\&$

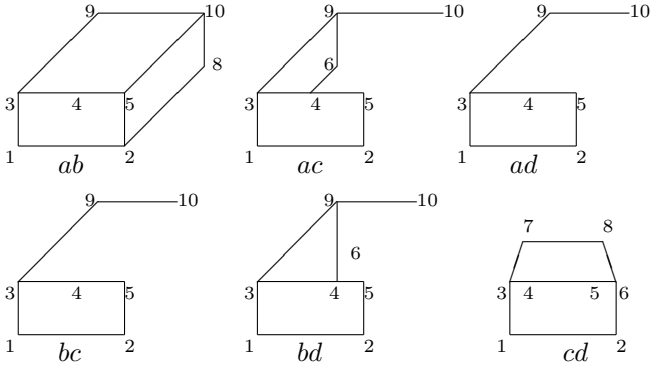


Рис. 4. Изображения, соответствующие выделенным подформулам

$V(x_3, x_9, x_4)\&V(x_3, x_9, x_5)\&V(x_3, x_9, x_1)\&V(x_5, x_2, x_4)\&$
$V(x_5, x_2, x_3)\&V(x_9, x_{10}, x_3)\&L(x_4, x_3, x_5)$, соответствующая этому изображению, определяет предикат 1-го уровня $p^1(x^1)$. Переменная первого уровня x^1 – это переменная для списка из 7 исходных переменных.

При этом, учитывая унификаторы, которые были определены при выделении всех общих (с точностью до имён переменных) подформул, при подстановке в формулы $A_i(\overline{x}_i)$ $(i = 1, ..., 4)$ значением переменной x^1 будут соответственно списки $(x_1, x_2, x_3, x_4, x_5, x_9, x_{10})$, $(x_1, x_2, x_4, x_5, x_6, x_7, x_8)$, $(x_1, x_2, x_3, x_4, x_6, x_7, x_8)$, $(x_1, x_2, x_3, x_4, x_6, x_9, x_{10})$.

После замены в формулах $A_i(\overline{x}_i)$ $(i = 1, ..., 4)$ подформул вида $P^1(\overline{y}^1)$ $(i = 1, ..., 4)$ с учётом полученных унификаторов на атомарную формулу $p^1(x^1)$ получим двухуровневое описание классов. Например, элементарная конъюнкция для изображения b на рисунке 2 примет вид $A_2^1(x^1, x_1, \ldots, x_8) =$

$$p^1(x^1)\&V(x_2, x_6, x_3)\&V(x_2, x_1, x_3)\&V(x_3, x_2, x_8)\&$$

$$V(x_5, x_4, x_7)\&V(x5, x7, x6)\&V(x6, x5, x8)\&V(x6, x4, x8)\&$$
$$V(x_6, x_8, x_2)\&V(x_7, x_5, x_4)\&V(x_7, x_8, x_5)\&V(x_7, x_8, x_4)\&$$
$$V(x_8, x_3, x_6)\&V(x_8, x_6, x_7)\&V(x_8, x_3, x_7).$$

Следует отметить, что несмотря на то, что формально количество переменных в формуле увеличилось, но при последующем

присвоении только переменной x^1 какого-то набора значений констант $(a_1, a_2, a_4, a_5, a_6, a_7, a_8)$, в формуле с неприсвоенными значениями останется только переменная x_3, так как остальным переменным будут присвоены значения, взятые из списка значений для переменной x^1.

Четыре признака второго уровня определяются элементарными конъюнкциями, соответствующими изображениям ab, ac, bd и cd на рисунке 4. При этом эти элементарные конъюнкции содержат признаки 1-го уровня и переменную 1-го уровня.

Элементарные конъюнкции $P_1^2(\overline{y}_1^2)$, $P_2^2(\overline{y}_2^2)$, $P_3^2(\overline{y}_3^2)$, $P_4^2(\overline{y}_4^2)$, соответствующие изображениям ab, ac, bd, cd на рисунке 4 и записанные с использованием предиката $p^1(x^1)$ определяют предикаты второго уровня $p_1^2(x_1^2)$, $p_2^2(x_2^2)$, $p_3^2(x_3^2)$, $p_4^2(x_4^2)$.

Например, подформула $P_1^2(\overline{y}_1^2) =$

$$p^1(x^1)\&V(x_2, x_5, x8)\&V(x_2, x_1, x_8)\&V(x_5, x_4, x_{10})\&V(x_5, x_3, x_{10})\&$$

$$V(x_8, x_2, x_{10})\&V(x_{10}, x_8, x_5)\&V(x_{10}, x_5, x_9)\&V(x_{10}, x_8, x_9)$$

соответствует изображению ab на рисунке 4. Здесь y_1^2 – переменная для списка, состоящего из переменных $(x^1, x_1, x_2, x_4, x_5, x_8, x_9, x_{10})$, причём переменная 1-го уровня x^1 – это переменная для списка исходных переменных $(x_1, x_2, x_3, x_4, x_5, x_9, x_{10})$. Заметим, что при определении значения для переменной 1-го уровня x^1 только переменная x_8 не получила значения.

После замены в формулах $A_i^1(\overline{x}_i^1)$ подформул $P_j^2(\overline{y}_i^2)$ $(i, j = 1, ..., 4)$ на атомарные формулы $p_j^2(x_j^2)$ получим двухуровневое описание классов. Например, элементарная конъюнкция для изображения a на рисунке 2 примет вид $A_1^2(x_1^2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10},) =$

$$p^1(x_1^2)\&V(x_4, x_3, x_6)\&V(x_4, x_6, x_5)\&V(x_6, x_4, x_9)\&$$

$$V(x_6, x_9, x_7)\&V(x_9, x_7, x_4)\&V(x_7, x_5, x_6)\&V(x_7, x_6, x_{10})\&$$

$$V(x_9, x_6, x_3)\&V(x_9, x_{10}, x_6)\&V(x_9, x_{10}, x_3)\&L(x_7, x_5, x_{10}).$$

Следует отметить, что все исходные переменные, кроме переменной x_6, входят в список переменных, определяющих x^2 .

Работа обучающего блока закончена. Сформирована 3-уровневая сеть. Схематически последовательность проверки 3-уровневой сетью истинности подформул, соответствующих фрагментам изображения, представлена на рисунке 5.

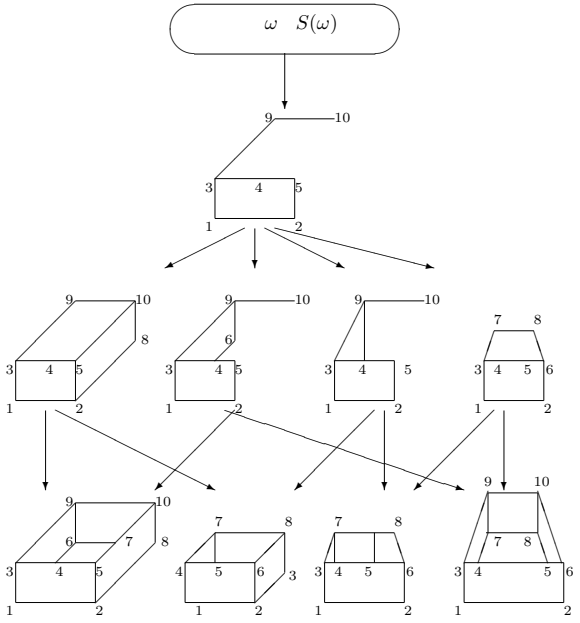


Рис. 5. Фрагменты изображений, выделяемые 3-уровневой сетью

В процессе распознавания все объекты, соответствующие изображениям на рисунке 2, будут распознаны правильно.

Пусть для распознавания представлен объект, изображённый на рисунке 6. Построенная сеть не сможет его распознать, так как формула, определяющая предикат 1-го уровня, не будет истинной на его описании.

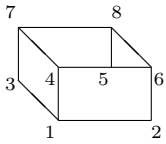


Рис. 6. Контрольное изображение

Добавим элементарную конъюнкцию, соответствующую его описанию, к ранее заданным исходным данным обучающего блока. По-

парная проверка неполной выводимости этой элементарной конъюнкции с уже имеющимися в обучающем блоке выделит новые подформулы. При этом предикат 1-го уровня будет задаваться элементарной конъюнкцией, соответствующей изображению на рисунке 7.

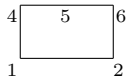


Рис. 7. Изображение, соответствующее новому предикату 1-го уровня

Предикаты 2-го уровня задаются формулами, соответствующими изображениям на рисунке 8.

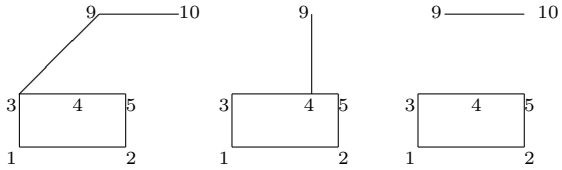


Рис. 8. Изображения, соответствующие новым предикатам 2-го уровня

Множество формул, определяющих предикаты 3-го уровня, совпадает с ранее построенным множеством предикатов 2-го уровня.

Таким образом, распознающий блок перестроен и представляет собой 4-уровневое описание класса. Схематически последовательность проверки 4-уровневой сетью истинности подформул, соответствующих фрагментам изображения, представлена на рисунке 9.

9. Заключение. В статье описан подход к формированию самоперестраивающейся нейронной сети с элементами, реализующими вычисление значения элементарной конъюнкции предикатных формул.

Основной проблемой при конструировании таких сетей в настоящий момент является детальная проработка способа хранения и передачи унификаторов, позволяющих отождествлять переменные и константы в выделенных подформулах и в формулах, в которые они подставляются.

Непроработанным также остаётся вопрос, какие же из выделенных подформул подставлять в исходные, если возможно несколько вариантов подстановок.

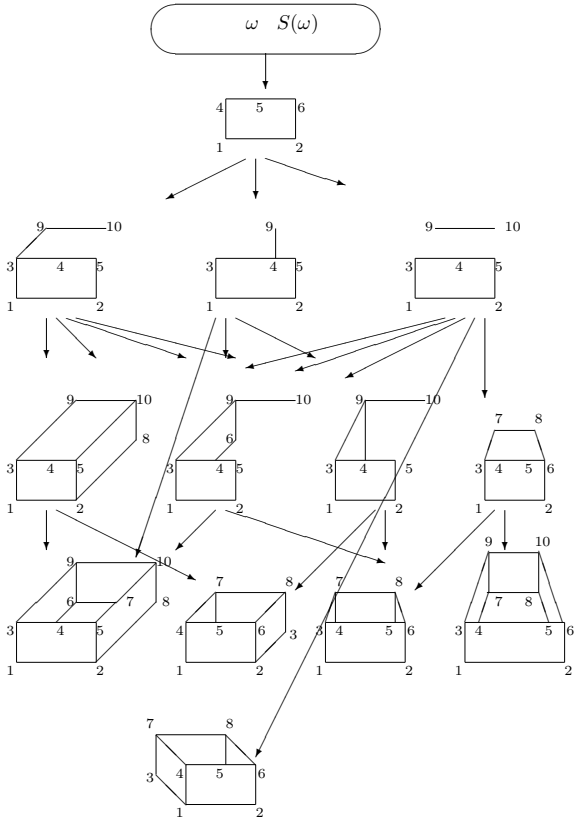


Рис. 9. Фрагменты изображений, выделяемые 4-уровневой сетью

Возможны варианты, когда выделенная подформула содержит несколько переменных более низкого уровня и, следовательно, задаёт отношение между ними. Интересны возможности использования взвешенных предикатов и переменных. По-видимому, эти веса должны меняться в зависимости от «верного» или «ошибочного» распознавания.

Список литературы

1. *Дуда Р., Харт П.* Распознавание образов и анализ сцен // М.: Мир, 1976. 511 с.
2. *Клини С.* Математическая логика // М.: Мир, 1973. 480 с.

3. *Комашинский В.И., Смирнов Д.А.* Нейронные сети и их применение в системах управления и связи // М.: Горячая линия–Телеком, 2002. 94 с.

4. *Косовская Т.М.* Доказательства оценок числа шагов решения некоторых задач распознавания образов, имеющих логические описания // Вестник Санкт-Петербургского университета. Сер. 1. 2007. Вып. 4. С. 82–90.

5. *Косовская Т.М., Тимофеев А.В.* Иерархическое описание классов и нейросетевое распознавание сложных образов // Нейрокомпьютеры: разработка, применение. 2007. № 6. С. 30 – 33.

6. *Косовская Т.М.* Многоуровневые описания классов для уменьшения числа шагов решения задач распознавания образов, описываемых формулами исчисления предикатов // Вестн. С.-Петербург.ун-та. Сер. 10. 2008. Вып. 1. С. 64 – 72.

7. *Косовская Т. М.* Частичная выводимость предикатных формул как средство распознавания объектов с неполной информацией // Вестн. С.-Петербург.ун-та. Сер. 10. 2009. Вып. 1. С. 74 – 84.

8. *Косовская Т.М.* Некоторые задачи искусственного интеллекта, допускающие формализацию на языке исчисления предикатов, и оценки числа шагов их решения // Труды СПИИРАН. 2010. Вып. 14. С. 58 – 75.

9. *Косовская Т.М.* Подход к решению задачи построения многоуровневого описания классов на языке исчисления предикатов // Труды СПИИРАН. 2014. № 3(34). С. 204 – 217.

10. *Рассел С., Норвиг П.* Искусственный интеллект: современный подход, 2-е изд. // Пер. с англ. М.: Издательский дом “Вильямс”, 2006. 1408 с.

References

1. Duda R.O., Hart P.E. Pattern Classification and Scene Analysis. A Wiley-Interscience Publication, John Wiley & Sons, New York London Sydney Toronto, 1973. 482 p. (Russ. ed.: Duda R., Hart P. Raspoznavanie obrazov i analiz scen. M.: Mir, 1976. 511p p.).

2. Kleene S.C. Mathematical logic. Dover Publications, New York: Wiley, 1967. (Russ. ed.: Klini S. Matematicheskaja logika. M.: Mir, 1973. 480 p.).

3. Komashinskiy V.I., Smirnov D.A. *Nejronnye seti i ih primenenie v sistemah upravleniya i svyazi* [Neural networks and their application in control systems and communication]. Moscow: Goryachaya liniya - Telekom, 2002. (In Russ.).

4. Kosovskaya T.M. [Proofs of the number of steps bounds for solving of some pattern recognition problems with logical description]. *Vestnik Sankt-Peterburgskogo Universiteta – Bulletin of St. Petersburg State University*. 2007, No. 4, pp. 82–90. (In Russ.).

5. Kosovskaya T.M., Timofeev A.V. [Hierarhical description of classes and neuron network recognition of complex patterns]. *Nejrokompjutery:*

razrabotka i primenenie – Neurocomputers: development, application. 2007. No. 6. pp. 30–33. (In Russ.)

6. Kosovskaya T.M. [Level descriptions of classes for decreasing step number of pattern recognition problem solving described by predicate calculus formulas]. *Vestnik Sankt-Peterburgskogo Universiteta – Bulletin of St. Petersburg State University*. 2008, vol. 10. No. 1, pp. 64–72. (In Russ.).
7. Kosovskaya T.M. [Partial hatchability predicate formulas as a means of recognition of objects with incomplete information]. *Vestnik Sankt-Peterburgskogo Universiteta – Bulletin of St. Petersburg State University*. 2009, vol. 10. No. 1, pp. 74–84. (In Russ.).
8. Kosovskaya T.M. [Some artificial intelligence problems permitting formalization by means of predicate calculus language and upper bounds of their solution steps]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2010. vol. 14, pp. 58 - 75. (In Russ.).
9. Kosovskaya T.M. [An approach to the construction of a level description of classes by means of a predicate calculus language]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2014. vol. 3(34), pp. 58–75. (In Russ.).
10. Russel S.J., Norvig P. Artificial Intelligence. A Modern Approach. Pearson Education. Inc. 2003. (Russ. ed.: Rassel S., Norvig P. Iskusstvennyj intellekt: sovremennyj podhod, 2-e izd. Per. s angl. M.: Izdatel'skij dom "Vil'jams", 2006. 1408 p.).

Косовская Татьяна Матвеевна — д-р физ.-мат. наук, доцент, профессор математико-механического факультета, Санкт-Петербургский государственный университет (СПбГУ), старший научный сотрудник, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: логический подход к решению задач искусственного интеллекта. Число научных публикаций — 85. kosovtm@gmail.com; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +79213232307.

Kosovskaya Tatiana Matveevna — Ph.D., Dr. Sci., associate professor, professor of computer science department, St.Petersburg State University (SPbSU), senior researcher of autonomous robotic systems laboratory, St. Petersburg Institute for Informatics and Automation of Russian Academy of Scientists (SPIIRAS). Research interests: logical approach to the solving of artificial intelligence problems, theory of complexity of algorithms. The number of publications — 85. kosovtm@gmail.com; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +79213232307.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проект № 14-08-01276-а).

Acknowledgements. This research is supported by RFBR (grant 14-08-01276-a).

РЕФЕРАТ

Косовская Т.М. **Самообучающаяся сеть с ячейками, реализующими предикатные формулы.**

Рассматривается модель перенастраиваемой сети с ячейками, реализующими предикатные формулы, имеющие вид элементарных конъюнкций. В отличие от классических нейронных сетей предлагаемая модель имеет два блока: блок обучения и блок решения.

При ошибках, возникающих при использовании блока решения, подключается блок обучения, который определяет дополнительные составные характеристики, присущие как ранее представленным объектам, так и вновь поступившему, для которого произошла ошибка. Кроме того, конфигурация сети не фиксируется заранее, а меняется каждый раз после работы блока обучения.

Рассматриваются задачи искусственного интеллекта, в которых исследуемый объект представлен как множество элементов, характеризующихся своими свойствами и отношениями между ними. Исследуемый объект задаётся своим описанием, представленным как множество постоянных атомарных формул (или их отрицаний), истинных для этого объекта. Целевые формулы записываются в виде дизъюнкции элементарных конъюнкций атомарных формул.

Поскольку задачи, допускающие такую формализацию, NP-трудны, то встаёт вопрос об уменьшении показателя степени экспоненты, ограничивающей число шагов решения задач. Построение уровневого описания целевых формул посредством выделения из них общих подформул позволяет существенно уменьшить число шагов алгоритмов, решающих рассматриваемые задачи. Такие уровневые описания соответствуют декомпозиции задачи большой размерности на несколько последовательно решаемых задач меньшей размерности. При этом в оценке числа шагов алгоритма вместо большого показателя степени экспоненты, соответствующего размерности исходной задачи, получается сумма, каждое слагаемое которой в показателе степени имеет размерность подзадачи.

В статье предлагается создание сети, в которой обучающий блок формирует многоуровневое описание классов. Работа обучающего блока заключается в создании уровневого описания классов на основе обучающей выборки посредством выделения общих составных характеристик всех элементов обучающей выборки. Описан алгоритм выделения таких составных характеристик, основанный на ранее введённом автором понятии неполной выводимости. Работа распознающего блока происходит в соответствии с приведённым в статье алгоритмом. Приведён модельный пример построения 3-уровневой сети и последующей её перестройки в 4-уровневую.

SUMMARY

Kosovskaya T.M. **Self-training Network with the Sells Implementing Predicate Formulas.**

A model of self-modificated predicate network with cells implementing predicate formulas in the form of elementary conjunction is suggested. Unlike a classical neuron network the proposed model has two blocks: a training block and a recognition block.

If a recognition block has a mistake then the control is transmitted to a training block. Always after a training block implementation the configuration of the recognition block is changed. The base of the proposed predicate network is logic-objective approach to AI problems solving and level description of classes.

Artificial Intelligence problems, an investigated object of which is presented as a set of elements characterized by their properties and relations between them, are under consideration. An investigated object is represented by its description as a set of constant atomic formulas (or their negations) which are true for this object. A goal formula is written in the form of disjunction of elementary conjunctions of atomic formulas.

As all these problems are NP-hard a problem of decreasing for the exponent of the number of steps upper bound for the decision algorithm is very important. Construction of a level description for goal formulas, by means of extracting their common sub-formulas, allows essentially to decrease the number of steps for an algorithm solving a problem under consideration. These descriptions correspond to decomposition of a high-dimensional problem up to several less-dimensional sub-problems which are solved sequentially. In this case the upper bound of number of steps of an algorithm is the sum of terms with exponents equal to dimensions of sub-problems instead of one term which exponent equals to the dimension of the initial problem.

A network with the training block creating a level description of classes is suggested in the paper. The training block implementation consists in the construction of a level description on the base of a training set be means of the extraction common complex characteristics of the training set elements. An algorithm of such an extraction based on the proposed earlier by the author notion of partial deducibility is described in the paper.

The recognition block implementation is made according to the algorithm described in the paper. A model example of 3-level network construction and further its reconstruction up to a 4-level network is presented.

УДК 004.04

С.Н. Баранов, В.В. Никифоров

ТРАНЗИТИВНОЕ НАСЛЕДОВАНИЕ ПРИОРИТЕТОВ В МНОГОЗАДАЧНЫХ ПРИЛОЖЕНИЯХ РЕАЛЬНОГО ВРЕМЕНИ

Баранов С.Н., Никифоров В.В. **Транзитивное наследование приоритетов в многозадачных приложениях реального времени.**

Аннотация. Рассматриваются методы контроля доступа задач к разделяемым ресурсам в программных приложениях для систем реального времени. Приводится детальное представление двух процедур наследования приоритетов задач: непосредственной и транзитивной. Сформулированы достаточные условия, при которых применение непосредственной процедуры предотвращает инверсию приоритетов. Предложена модификация транзитивной процедуры снимающая известные ограничения на структуру приложения, накладываемые ее традиционной реализацией. Эта модификация, кроме того, обеспечивает динамическое обнаружение некорректных ситуаций типа взаимного блокирования задач с возможностью запланированной реакции на такие ситуации.

Ключевые слова: системы реального времени, модели многозадачных приложений, выполнимость задач, протоколы доступа к разделяемым ресурсам.

Baranov S.N., Nikiforov V.V. **Transitive Priority Inheritance in Real-Time Multi-Task Applications.**

Abstract. Control procedures for accessing shared resources in multi-task real-time software applications are analyzed. Two approaches to preventing priority inversion – direct and transitive procedures of priority inheritance – are analyzed in detail. To illustrate the considered notions and statements, concrete examples of multi-task application configurations are provided along with their execution diagrams under particular scenarios of system events, which demonstrate insufficient response time of tasks with sufficient computing resources and even mutual task clinches. The nomenclature of attributes to be included into task and resource descriptors for task management with the two priority inheritance procedures is proposed. The sufficient conditions for the direct procedure to prevent priority inheritance are formulated. The procedure of transitive priority inheritance is demonstrated to be capable of detecting a mutual task clinch in the application.

Keywords: real-time systems, multi-task application models, task feasibility, shared resources access protocols.

1. Введение. Программные продукты, предоставляющие пользователю широкий спектр функциональных возможностей, обычно строятся в виде многозадачных приложений, состоящих из ряда задач τ_1, τ_2, …, τ_n, каждой из которых предоставляется доступ к различным активным (исполнительным) и пассивным (информационным) ресурсам как локальным – доступным лишь одной задаче, так и глобальным – попеременно доступным разным задачам. Такое построение программных продуктов типично для систем реального времени (СРВ) [1], встроенных систем, систем имитационного моделирования и систем, ориентированных на исполнение с использованием многопроцессорных компьютеров и многоядерных процессоров [2].

При создании многозадачных программных приложений обычно возникают две проблемы:

1) определение порядка предоставления задачам ресурса процессора (процессорного времени), гарантирующего своевременность их исполнения [3];

2) обеспечение целостности глобальных информационных ресурсов (т.е., предотвращение одновременного доступа к нему двух и более заданий) [4].

Первая проблема решается применением дисциплины планирования, обеспечивающей выполнимость задач – гарантированную своевременность их выполнения. Эффективные дисциплины планирования для однопроцессорных систем с классическими одноядерными процессорами предлагались с начала 1970-х годов – в частности RM (Rate-Monotonic) и EDF (Earliest Deadline First) дисциплины планирования [5]. В работе [6] было показано, что RM и EDF могут терять свою эффективность при использовании многопроцессорных систем. В 2000-е годы предлагались модификации этих дисциплин планирования, обеспечивающие приемлемую эффективность для систем на многоядерных процессорах [7]. В настоящее время как отечественными (например, [8]), так и зарубежными (например, [9], [10]) исследователями продолжается поиск эффективных дисциплин планирования. Специально для многоядерных процессоров разработаны высокоэффективные Pfair (справедливо-пропорциональные) дисциплины планирования с квантованием [11] и без квантования [12] интервалов выделяемого процессорного времени. Вторая проблема решается путем использования протоколов доступа к разделяемым информационным ресурсам [13].

2. Структура и параметры задач. Будем считать каждую задачу последовательной программой, замкнутой в себе по передачам управления. Задачи активизируются в порядке реакции на внешние события через какие-то промежутки времени, не меньшие некоторой величины – их периода. Таким образом, каждая задача τ_i характеризуется своим периодом T_i, весом C_i – объемом процессорного времени, необходимого для выполнения ее вычислительной работы, предельным сроком D_i для завершения ее исполнения, и фазой P_i, задающей момент первой активизации данной задачи. Очевидно, что для всех задач должно выполняться неравенство $C_i \le D_i$.

Очередная (k-ая) активизация задачи τ_i означает порождение задания ${}_k\tau_i$ (порождение очередного k-ого экземпляра задачи τ_i). Задание является активным в некоторый конкретный момент времени, если к этому моменту оно уже порождено, но еще не завершено. То есть, ка-

ждое задание является активным в рамках интервала его существования – от момента его порождения до момента его завершения. Порождение задания означает увеличение числа претендентов на ресурс процессора (процессорное время) и, следовательно, изменение условий его распределения. Поскольку изменение условий распределения системных ресурсов в общем случае является следствием какого-либо системного события, то порождение задания следует рассматривать как его разновидность. Другим примером системного события является завершение задания, приводящее к уменьшению числа претендентов на ресурс процессора.

Участок кода задачи τ_i, в рамках которого осуществляется доступ к какому-либо из глобальных информационных ресурсов g, называется критическим интервалом по доступу к этому ресурсу [14]. Для обеспечения целостности глобального информационного ресурса g_x используются механизмы контроля доступа к критическим интервалам, которые строятся на базе синхронизирующих элементов – мьютексов [16]. Для каждого разделяемого информационного ресурса g_x формируется мьютекс *m_x*, принимающий одно из двух значений: «открыт», если ресурс свободен, и «закрыт», если ресурс занят. Критический интервал по доступу к ресурсу g_x обрамляется специальными системными операторами над соответствующим мьютексом *m_x*. На входе в критический интервал выполняется системный оператор запроса ресурса g_x *lock(m_x)*, переводящий мьютекс *m_x* из состояния «открыт» в состояние "закрыт", на выходе – оператор освобождения ресурса g_x *unlock(m_x)*, переводящий этот мьютекс из состояния «закрыт» в состояние «открыт». Исполнение задания, запрашивающего занятый ресурс, приостанавливается до момента его освобождения. Различные критические интервалы по доступу к одному и тому же ресурсу g_x называются однотипными критическими интервалами.

Операторы *lock* предполагают системное событие, поскольку приводят либо к занятию глобального информационного ресурса, либо (если ресурс уже занят) к приостановке текущего задания с соответствующим уменьшением числа претендентов на ресурс процессора. Аналогично, операторы *unlock* также вызывают системное событие, поскольку приводят к освобождению глобального информационного ресурса и в случае возобновления ожидающего его задания – увеличение числа претендентов на процессорное время.

Структура задачи представляется в виде конечной последовательности сегментов, где каждый сегмент выполняет некоторое вычисление в течение некоторого отрезка процессорного времени (длина сегмента) и завершается одним из следующих системных событий:

«занять ресурс», «освободить ресурс», либо «завершить задачу». Вес задачи равен сумме длин ее сегментов, поскольку предполагается, что продолжительность системных событий, завершающих каждый сегмент, пренебрежительно мала.

Порядок предоставления задачам процессорного времени определяется используемой дисциплиной планирования. Любую дисциплину планирования можно выразить в виде способа наделения заданий целочисленными приоритетами, по которому активным задачам присваиваются определенные значения приоритетов, а ресурс процессора предоставляется наиболее приоритетной из активных задач. В большинстве реализаций многозадачных приложений приоритеты задач назначаются статически. Условимся индексировать задачи τ_1, τ_2, …, τ_n, составляющие приложение, в порядке снижения их приоритета: τ_1 – наиболее приоритетная задача, τ_n – наименее приоритетная, так что номер задачи задает ее приоритет.

Таким образом, в каждый момент времени своего существования активное задание ${}_k\tau_i$ может находиться в одном из трех состояний:

• использует ресурс процессора (является текущим заданием);

• ожидает освобождения процессора более приоритетным заданием;

• ожидает освобождения ресурса, занятого другим активным заданием.

В общем случае при исполнении многозадачных приложений, в которых целостность глобальных информационных ресурсов обеспечивается механизмами защиты одновременного попадания различных задач в однотипные критические интервалы, возможно возникновение таких некорректных ситуаций, как инверсия приоритетов и взаимное блокирование задач [16]. В первом случае менее приоритетное задание занимает ресурс процессора, тогда как более приоритетное задание ожидает освобождения информационного ресурса, занятого этим или другим менее приоритетным заданием; во втором – каждое из блокирующих заданий ожидает освобождения информационного ресурса, занятого другим заданием. Для предотвращения инверсии приоритетов предлагается дополнить дисциплину планирования (способ назначения приоритетов активным заданиям) функцией наследования приоритетов. Ниже описаны два варианта реализации этой функции – упрощенное непосредственное и более сложное транзитивное наследование приоритетов. Сформулированы условия, при которых использование непосредственного наследования приоритетов гарантирует предотвращение инверсии приоритетов. Представлен также способ такой модификации транзитивного наследования приоритетов, который позво-

ляет динамически обнаруживать возникновение ситуаций типа взаимного блокирования задач с возможностью запланированной реакции на такую ситуацию (выход из нее).

3. Инверсия приоритетов. Представляемые ниже особенности реализации функции наследования приоритетов иллюстрируются диаграммами исполнения многозадачного приложения из четырех задач τ_1, τ_2, τ_3 и τ_4, разделяющих два ресурса g_1 и g_2. Структура задач изображена на рис.1 в соответствии с предложенным в [17] подходом к представлению межзадачных интерфейсов средствами языка XML.

Задача τ_1

```
<task name="t_1" prio="1"
                 period="15">
  <segment length="1"
           interface="m_1"
           op_type="lock"/>
  <segment length="1"
           interface="m_1"
           op_type="unlock"/>
  <segment length="1"
           op_type="end"/>
</task>
```

Задача τ_2

```
<task name="task_1" prio="2"
                    period="35">
  <segment length="9"
           op_type="end"/>
</task>
```

Задача τ_3

```
<task name="t_3" prio="3"
                 period="25">
  <segment length="1"
           interface="m_1"
           op_type="lock"/>
  <segment length="2"
           interface="m_2"
           op_type="lock"/>
  <segment length="1"
           interface="m_2"
           op_type="unlock"/>
  <segment length="1"
           interface="m_1"
           op_type="unlock"/>
  <segment length="1"
           op_type="end"/>
</task>
```

Задача τ_4

```
<task name="t_1" prio="1">
                 period="45">
  <segment length="2"
           interface="m_2"
           op_type="lock"/>
  <segment length="4"
           interface="m_2"
           op_type="unlock"/>
  <segment length="1"
           op_type="end"/>
</task>
```

Рис. 1. Вариант приложения из четырех задач, разделяющих два ресурса

Имеющая самый высокий приоритет задача τ_1 использует ресурс g_1, доступ к которому контролируется мьютексом *m_1*. Код задачи τ_1 состоит из трех сегментов, требующих для своего исполнения по одной единице процессорного времени. Первый сегмент заканчивается операцией *lock(m_1)* запроса доступа к ресурсу g_1. Второй сегмент – критический интервал по доступу к g_1 – заканчивается операцией *unlock(m_1)* освобождения этого ресурса. Заключительный третий сегмент заканчивается операцией завершения задачи.

Код задачи τ_2 состоит из единственного сегмента, который заканчивается оператором завершения задачи и требует для своего исполнения 9 единиц процессорного времени. Значение 2 ее приоритета

означает, что задача τ_2 менее приоритетна, чем задача τ_1, но более приоритетна, чем τ_3 и τ_4.

Код задачи τ_3 содержит 5 сегментов. Первый заканчивается операцией *lock(m_1)* запроса доступа к ресурсу g_1. Критический интервал по доступу задачи τ_3 к ресурсу g_1 включает второй, третий и четвертый сегменты ее кода. Второй сегмент заканчивается запросом доступа к ресурсу g_2. В рамках третьего сегмента, который завершается освобождением ресурса g_2, задача τ_3 владеет как ресурсом g_1, так и ресурсом g_2. Четвертый сегмент завершает критический интервал по доступу к g_1, а пятый сегмент завершает исполнение задачи τ_3.

Код наименее приоритетной задачи τ_4 содержит 3 сегмента: начальный сегмент длиной 2; второй сегмент, являющийся критическим интервалом по доступу задачи τ_4 к ресурсу g_2 длиной 4, и заключительный сегмент длиной 1.

Периоды T_1, T_2, ${}_1T_3$ и T_4 активизации задач равны соответственно 15, 35, 25 и 45 единицам времени, а их фазы равны 5, 5, 3 и 0 единиц времени соответственно. Условимся считать, что для каждой из задач τ_i предельный срок выполнения равен ее периоду: $D_i = T_i$.

Состав действий, реализуемых при исполнении операций *lock* и *unlock*, определяется выбором протокола доступа к глобальным информационным ресурсам. Простейший протокол, обеспечивающий сохранение целостности каждого из ресурсов g_x, предписывает формирование для каждого ресурса g_x синхронизирующего элемента типа mutex с полями состояние («открыт» или «закрыт») и список заданий, ожидающих освобождения данного ресурса.

При инициализации мьютекса, соответствующего ресурсу g, его поле состояния получает значение «открыт» (т.е., ресурс свободен), а список заданий, ожидающих освобождения ресурса, делается пустым.

В реализации простейшего протокола выполнение операций *lock* и *unlock* сводится к действиям, приведенным в таблице 1. При непустом списке заданий, ожидающих освобождения ресурса, текущее задание, выполняющее операцию *unlock(m_x)*, может быть вытеснено более приоритетным новым владельцем ресурса g_x, стоявшим в начале этого списка.

На рисунке 2*а* приведена диаграмма исполнения приложения из четырех задач с параметрами, соответствующими рисунку 1, при использовании протокола доступа, представленного в таблице 1. Диаграмма соответствует конкретному сценарию системных событий (конкретному порядку активизации задач) при котором возникает инверсия приоритетов, приводящая к задержке исполнения высокоприоритетных заданий.

Таблица 1. Реализация простейшего протокола доступа к глобальным ресурсам

Операция	Условие	Изменения состояний задач и ресурсов
lock(m)	Мьютекс *m*, соответствующий ресурсу *g*, открыт	Мьютекс `m` переводится в состояние "закрыт". Задание, выполнившее операцию *lock(m)*, становится владельцем ресурса *g*, приступает к исполнению критического интервала по этому ресурсу.
	Мьютекс *m*, соответствующий ресурсу *g*, закрыт	Дескриптор задания, выполняющего операцию *lock(m)*, переносится из списка претендентов на ресурс процессора в список заданий, ожидающих освобождения ресурса *g*
unlock(m)	Список заданий, ожидающих освобождения ресурса *g*, пуст	Мьютекс `m` переводится в состояние «открыт». Задание выполнившее операцию *unlock(m)*, продолжает исполняться
	Список заданий, ожидающих освобождения ресурса *g*, не пуст	Задание, возглавляющее список ожидающих освобождения ресурса *g*, становится владельцем ресурса *g*, соответствующего мьютексу *m*; его дескриптор переносится из этого списка в список претендентов на ресурс процессора

В таблице 2 приведен комментарий к диаграмме на рисунке 2*а*: перечень действий, выполняемых в моменты возникновения каждого из системных событий. Для упрощения изложения вторая активизация задачи τ_1 (т.е. порождение задания ${}_2\tau_1$ в момент времени после t=20) не рассматривается.

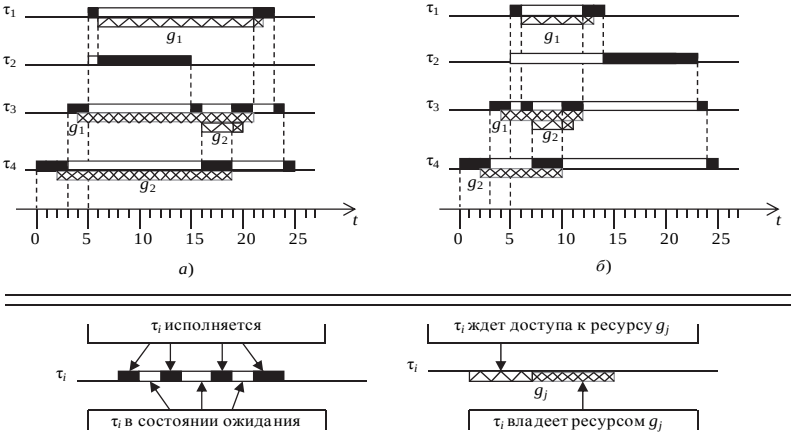


Рис. 2. Исполнение комплекса задач, использующих разделяемые ресурсы

Таблица 2. Сценарий системных событий, вызывающий инверсию приоритетов

Время	Событие	Изменения состояний заданий и ресурсов
t=0	Активизация задачи τ_4	Порождается и инициализируется задание ${}_1\tau_4$ типа τ_4, процессор переключается на его исполнение.
t=2	Запрос ресурса g_2 заданием ${}_1\tau_4$	Задание ${}_1\tau_4$ запрашивает доступ к свободному ресурсу g_2. Ресурс g_2 предоставляется в его распоряжение и задание ${}_1\tau_4$ приступает к исполнению своего критического интервала по доступу к нему.
t=3	Активизация задачи τ_3	Порождается задание ${}_1\tau_3$ типа τ_3. Приоритет τ_3 выше, чем у τ_4, вследствие чего задание ${}_1\tau_3$ инициализируется и вытесняет задание ${}_1\tau_4$ – процессор переключается на исполнение ${}_1\tau_3$.
t=4	Запрос ресурса g_1 заданием ${}_1\tau_3$	Задание ${}_1\tau_3$ запрашивает доступ к свободному ресурсу g_1. Ресурс g_1 ему предоставляется, и процессор продолжает исполнение ${}_1\tau_3$ уже в рамках критического интервала по доступу к g_1.
t=5	Активизация задач τ_1 и τ_2	Порождаются задания ${}_1\tau_1$ типа τ_1 и ${}_1\tau_2$ типа τ_2. Приоритет τ_1 выше, чем приоритет текущего задания ${}_1\tau_3$, – выполняется инициализация задания ${}_1\tau_1$, оно вытесняет ${}_1\tau_3$, процессор переключается на исполнение ${}_1\tau_1$.
t=6	Запрос ресурса g_1 заданием ${}_1\tau_1$	Задание ${}_1\tau_1$, запрашивает доступ к ресурсу g_1. Поскольку ресурс g_1 занят заданием ${}_1\tau_3$, задание ${}_1\tau_1$ переводится в состояние ожидания момента освобождения g_1. Освободившийся таким образом процессор переключается на исполнение наиболее приоритетного из заданий, готовых использовать ресурс процессора. Таким заданием является задание ${}_1\tau_2$ – оно инициализируется, и процессор приступает к его исполнению.
t=15	Завершение задания ${}_1\tau_2$	В результате завершения задания ${}_1\tau_2$ процессор освобождается и переключается на продолжение исполнения задания ${}_1\tau_3$, а менее приоритетное задание ${}_1\tau_4$ остается в состоянии ожидания момента предоставления ему процессора.
t=16	Запрос ресурса g_2 заданием ${}_1\tau_3$	Ресурс g_2 занят заданием ${}_1\tau_4$ – исполнение задания ${}_1\tau_3$ приостанавливается, процессор переключается на исполнение задания ${}_1\tau_4$.
t=19	Задание ${}_1\tau_4$ освобождает ресурс g_2	Завершается исполнение критического интервала задания ${}_1\tau_4$ по ресурсу g_2. Освободившийся ресурс g_2 предоставляется ожидающему его заданию ${}_1\tau_3$. Процессор переключается с задания ${}_1\tau_4$ на исполнение более приоритетного задания ${}_1\tau_3$.
t=20	Задание ${}_1\tau_3$ освобождает ресурс g_2	Завершается исполнение пересечения критических интервалов задания ${}_1\tau_3$ по ресурсам g_1 и g_2. Список заданий, претендующих на ресурс процессора, не изменяется. Процессор остается в распоряжении задания ${}_1\tau_3$.
t=21	Задание ${}_1\tau_3$ освобождает ресурс g_1	Освободившийся ресурс g_1 предоставляется ожидающему его более приоритетному заданию ${}_1\tau_1$. Процессор переключается на исполнение задания ${}_1\tau_1$.
t=23	Завершение задания ${}_1\tau_1$	Процессор переключается на исполнение завершающих участков кода задания ${}_1\tau_3$ и затем задания ${}_1\tau_4$.

Максимальная продолжительность исполнения заданий типа τ_i обозначается символом R_i и называется временем отклика задачи τ_i. Требование своевременности исполнения отдельной задачи τ_i, входящей в состав приложения, выражается неравенством $R_i \le D_i$.

По рисунку 2*а* видно: продолжительность интервала существования задания ${}_1\tau_1$, порождаемого в момент t=5 и заканчивающегося в момент t=23, равна 18, что превышает заданный предельный срок D_1=15. Следовательно, для приложения на рисунке 1 при использовании протокола согласно таблице 1 требование $R_i \le D_i$ не выполняется.

Предельная продолжительность исполнения задачи τ_1 нарушается, несмотря на то, что она является самой приоритетной из всех задач и для ее исполнения требуется лишь 3 единицы процессорного времени, тогда как остальные 15 единиц процессорного времени на интервале (5, 23) расходуются менее приоритетными заданиями ${}_1\tau_2$, ${}_1\tau_3$ и ${}_1\tau_4$. Из этих 15-ти единиц 6 расходуются блокирующими задачами τ_3 и τ_4. В течение 3-х единиц времени (15, 19–20) задание ${}_1\tau_3$ блокирует задание ${}_1\tau_1$ непосредственно, занимая ресурс g_1, требующийся заданию ${}_1\tau_1$. Еще три единицы времени (16-18) расходуются заданием ${}_1\tau_4$ для освобождения ресурса g_2, без доступа к которому задание ${}_1\tau_3$ не может завершить критический интервал по ресурсу g_1. Но в рамках используемого протокола (таблица 1) большая часть интервала (5, 23) расходуется на исполнение задания ${}_1\tau_2$, несмотря на то, что у него и приоритет ниже, чем у τ_1, и сама задача τ_2 не связана ни прямо, ни опосредованно с требуемым задачей τ_1 ресурсом. При добавлении в приложение других среднеприоритетных задач (задач со значениями приоритетов ниже, чем у задачи τ_1, но выше, чем у задач τ_3 и τ_4) их исполнение может еще более тормозить исполнение самой приоритетной задачи τ_1.

Для обозначения рассмотренного влияния среднеприоритетных задач на увеличение времени отклика высокоприоритетной задачи, разделяющей глобальные ресурсы с низкоприоритетными задачами, используется термин *инверсия приоритетов*. Предотвращение инверсии приоритетов достигается путем дополнения протокола доступа задач к глобальным ресурсам функцией наследования приоритетов.

4. Непосредственное наследование приоритетов. Отличительной особенностью протокола доступа к ресурсам с наследованием приоритетов является возможность изменения приоритета заданий в процессе их исполнения. При этом различаются приоритет инициализации (исходный приоритет) задания и его действующий (текущий) приоритет. Приоритет инициализации равен приоритету задачи, из которой порождается данное задание; он один и тот же для всех таких заданий. Действующий приоритет устанавливается равным приоритету инициализации задания в момент его порождения и в дальнейшем может изменяться, в частности, через механизм наследования приоритетов.

Принцип наследования приоритетов выражается следующим образом: если при выполнении операции *lock(m_x)* высокоприоритетным заданием ${}_m\tau_i$ ресурс g_x, соответствующий мьютексу *m_x*, занят менее приоритетным заданием ${}_n\tau_j$, то действующий приоритет задания ${}_n\tau_j$ временно, до освобождения им ресурса g_x, устанавливается равным действующему приоритету задания ${}_m\tau_i$.

Для реализации наследования приоритетов структуру для дескрипторов заданий следует дополнить полями для приоритета инициализации и действующего приоритета, а структуру мьютекса – полем для указателя на дескриптор задания, владеющего данным ресурсом в настоящий момент.

При дополнении протокола доступа к разделяемым ресурсам функцией наследования приоритетов процедуры, приведенные в таблице 1, должны быть дополнены действиями для предотвращения инверсии приоритетов. В таблице 3 приведен их перечень для непосредственного наследования приоритетов.

Таблица 3. Непосредственное наследование приоритетов

Операция	Условие	Действия, выполняемые в добавление к приведенным в таблице 1
lock(m)	Мьютекс *m*, соответствующий ресурсу *g*, открыт	Указатель на дескриптор текущего задания ${}_x\tau_i$ сохраняется в структуре мьютекса.
	Мьютекс *m*, соответствующий ресурсу *g*, закрыт	Действующий приоритет задания ${}_y\tau_j$, владеющего ресурсом *g*, повышается до значения действующего приоритета текущего задания
unlock(m)	Задание ${}_y\tau_j$, выполняющее операцию *unlock(m)*, владеет и другими ресурсами с непустыми списками ожидающих заданий	В этих списках ищется задание ${}_x\tau_i$ с максимальным значением действующего приоритета. Приоритет задания ${}_y\tau_j$, освобождающего ресурс, делается равным высшему из двух приоритетов: – действующий приоритет ${}_x\tau_i$, – приоритет инициализации ${}_y\tau_j$. Указатель на дескриптор задания – нового владельца данного ресурса сохраняется в структуре мьютекса.
	Таких ресурсов нет	Действующий приоритет текущего задания делается равным его приоритету инициализации. Указатель на дескриптор нового владельца данного ресурса сохраняется в структуре мьютекса.

На рисунке 2*б* приведена диаграмма исполнения примера приложения (рисунок 1) при использовании протокола доступа (таблица 1), дополненного реализацией непосредственного наследования приоритетов (таблица 3). Диаграмма соответствует тому же сценарию

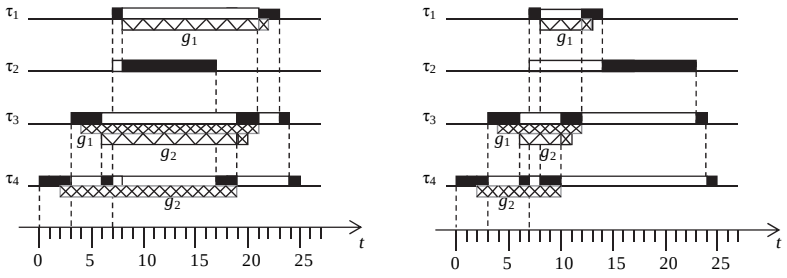
системных событий, что и в диаграмме рисунок 2*а*, но благодаря дополнительным действиям (таблица 3), исполнение задания ${}_1\tau_2$ выносится за рамки интервала существования задания ${}_1\tau_1$. В результате нарушение предельного срока исполнения задания ${}_1\tau_1$ не происходит.

Предлагаемые в классической работе [16] действия по выполнению операции *unlock* ориентированы на соблюдение следующего структурного ограничения: допускаются только вложенные критические интервалы (если два критических интервала пересекаются, то один из них должен быть вложен в другой). Такое требование вложенности критических интервалов, с одной стороны, позволяет существенно упростить реализацию процедуры *unlock*, но, с другой стороны, сужает возможности программистов, вынужденных учитывать приведенное ограничение. Модификация процедуры *unlock*, приведенная в таблице 3, пригодна для обслуживания приложений с любыми вариантами пересечений критических интервалов.

5. Отношения предшествования и транзитивное наследование приоритетов. В ходе отраженного на рисунке 2*б* исполнения приложения, соответствующего рисунку 1, при запросе в момент времени t=6 ресурса g_1 заданием ${}_1\tau_1$ обнаруживается, что требуемый ресурс занят заданием ${}_1\tau_3$. Так возникает отношение предшествования ${}_1\tau_1 \succ {}_1\tau_3$: прежде, чем сможет продолжиться исполнение задания ${}_1\tau_1$, задание ${}_1\tau_3$ должно завершить исполнение критического интервала по ресурсу g_1. В соответствии с таблицей 3, в момент t=6 действующий приоритет задания ${}_1\tau_3$ устанавливается равным действующему приоритету задания ${}_1\tau_1$. Далее, в момент времени t=7 аналогичным образом возникает отношение предшествования ${}_1\tau_3 \succ {}_1\tau_4$, действующий приоритет задания ${}_1\tau_4$ устанавливается равным действующему приоритету ${}_1\tau_3$. Поскольку к моменту t=7 еще сохраняется отношение ${}_1\tau_1 \succ {}_1\tau_3$, возникает цепочка ${}_1\tau_1 \succ {}_1\tau_3 \succ {}_1\tau_4$ отношений предшествования. При сценарии системных событий, отраженном на рисунке 2, по этой цепочке в соответствии с процедурой непосредственного наследования приоритетов действующий приоритет задания ${}_1\tau_1$ передается (транзитивно, через ${}_1\tau_3$) заданию ${}_1\tau_4$. Для предотвращения инверсии приоритетов процедура наследования должна при любом сценарии обеспечивать транзитивную передачу действующего приоритета через всю цепочку отношений предшествования [18].

Для приложения, представленного на рисунке 1, возможны такие сценарии системных событий, при которых непосредственное наследование приоритетов не обеспечивает передачу действующего приоритета заблокированного задания по всей цепочке отношений предшествования. Пример приведен на рисунке 3.

Сценарий системных событий, отражаемых диаграммами на рисунках 3*в* и 3*б*, отличается от сценария на рисунках 2*а* и 2*б* только тем, что задачи τ_1 и τ_2 активизируются в момент t=7 вместо t=5. В этих условиях непосредственное наследование приоритетов не предотвращает инверсии приоритетов (рисунок 3*а*).



а) б)

Рис. 3. Пример реализации процедур наследования приоритетов: а) непосредственное наследование приоритетов; б) транзитивное наследование приоритетов

К моменту порождения задания ${}_1\tau_1$ уже возникло (в момент t=6) отношение предшествования ${}_1\tau_3 \succ {}_1\tau_4$, и в соответствии с таблицей 3, действующий приоритет задания ${}_1\tau_4$ был повышен до приоритета инициализации ${}_1\tau_3$. Далее, в момент t=8 добавляется отношение ${}_1\tau_1 \succ {}_1\tau_3$ и возникает цепочка ${}_1\tau_1 \succ {}_1\tau_3 \succ {}_1\tau_4$. Однако непосредственное наследование приоритетов не обеспечивает транзитивную передачу действующего приоритета через всю эту цепочку от задания ${}_1\tau_1$ к заданию ${}_1\tau_4$. Действительно, в момент t=8 в соответствии со второй строкой таблицы 3 действующий приоритет задания ${}_1\tau_3$, владеющего ресурсом g_1, становится равным наивысшему приоритету – приоритету инициализации задания ${}_1\tau_1$. Но действующий приоритет текущего задания ${}_1\tau_2$ остается равным его приоритету инициализации. Поэтому процессор переключается на исполнение задания ${}_1\tau_2$, чей приоритет выше действующего приоритета задания ${}_1\tau_4$. По этой причине непосредственное наследование приоритетов, как показано на рисунке 3*а*, не предотвращает возникновения инверсии приоритетов – оно не обеспечивает передачу действующего приоритета ${}_1\tau_1$ через всю возникшую цепочку отношений предшествования.

6. Транзитивная процедура наследования приоритетов. Для предотвращения инверсии приоритетов при любом сценарии системных событий процедуру непосредственного наследования приоритетов необходимо дополнить действиями, обеспечивающими транзитивную передачу действующего приоритета блокируемого задания через всю

цепочку отношений предшествования. Ниже приведены дополнения в процедуру наследования приоритетов, обеспечивающие такую транзитивную передачу. Получающуюся в результате процедуру назовем транзитивной процедурой наследования приоритетов. Отметим, что в любой момент времени функционирования многозадачного приложения количество отношений предшествования ограничено числом активных заданий.

Для выполнения транзитивной процедуры следует дополнить дескриптор задания полем для номера ресурса, блокирующего исполнение этого задания.

При порождении задания в этом поле помещается значение, свидетельствующее о том, что задание не находится в состоянии ожидания глобального ресурса. В дальнейшем, при переходе задания в состояние ожидания перед входом в критический интервал, охраняемый мьютексом, в это поле заносится номер блокирующего ресурса.

Дополнения, переводящие процедуру непосредственного наследования приоритетов в транзитивную, касаются действий, выполняемых в рамках операции *lock(m_x)*, если мьютекс *m_x* закрыт.

При выполнении текущим заданием операции *lock(m_x)* вслед за действиями, приведенными в таблицах 1 и 3, в дескриптор этого задания помещается номер *m_x* блокирующего ресурса. После этого необходимо обработать всю цепочку начинающихся с этого задания отношений предшествования, чтобы повысить действующий приоритет каждого из них.

Обработка цепочки отношений предшествования состоит в циклическом выполнении приведенной ниже последовательности шагов (шаги 1-3). В их описании используется символ τ^*. В рамках первой итерации цикла символ τ^* означает задание, владеющее ресурсом и указатель на дескриптор которого сохранен в структуре мьютекса. В рамках каждой последующей итерации τ^* означает очередное задание из цепочки отношений предшествования.

Шаг 1. Перед входом в этот шаг действующий приоритет задания τ^* уже повышен. Проверяется поле дескриптора задания τ^* с номером ресурса, освобождения которого оно ожидает. Его нулевое значение говорит о том, что задание τ^* не блокировано и находится в списке заданий, претендующих на использование процессора. Отсюда следует, что обработка цепочки отношений предшествования завершена – в цепочке не осталось заданий, которым необходимо повысить действующий приоритет. Выполняется выход из цикла.

Шаг 2 (выполняется при ненулевом значении проверенного поля). Выявляется задание, требующее (транзитом) повышения своего

действующего приоритета. Таковым является задание, владеющее ресурсом с номером из проверенного поля. Символ τ* с этого момента означает новое звено цепочки отношений предшествования – найденное задание.

Шаг 3. Действующий приоритет задания τ* повышается до значения действующего приоритета текущего задания, выполняющего операцию *lock(m_x)*. Выполняется переход к шагу 1 – к очередной итерации цикла по обработке цепочки отношений предшествования.

Результат применения такой транзитивной процедуры наследования приоритетов представлен на рисунке 3*б*. Инверсия приоритетов предотвращена.

7. Влияние инверсии приоритетов на выполнимость задач. Ключевым требованием к приложению реального времени является обеспечение своевременности выполнения составляющих его задач. Задача τ_i называется выполнимой, если максимально возможная продолжительность ее исполнения (время отклика R_i) не превосходит заданного значения предельного срока D_i: $R_i \le D_i$. Приложение называется выполнимым, если гарантируется выполнимость каждой из составляющих его задач.

Возможность возникновения инверсии приоритетов в системе, не оснащенной механизмом наследования приоритетов, не обязательно приводит к нарушению выполнимости. Оценка выполнимости программных приложений СРВ осуществляется известными методами на этапе проектирования по заданной конфигурации приложения [14].

Для приложений, допускающих инверсию приоритетов, учет этого обстоятельства может быть включен в методы оценки выполнимости. Механизм наследования приоритетов становится излишним, если оценки R_i времени отклика задач, увеличенные вследствие инверсии приоритетов, остаются в рамках заданных предельных сроков D_i.

Кроме того, инверсия приоритетов возможна не во всяком многозадачном приложении с разделяемыми ресурсами. Для ее возникновения необходимо наличие среднеприоритетных задач между двумя разделяющими общий ресурс высокоприоритетной и низкоприоритетной задачами.

При работе приложений, не содержащих задач с пересекающимися критическими интервалами, нет необходимости включения механизмов транзитивного наследования приоритетов в управление заданиями. Наличие таких пересечений является необходимым условием возникновения цепочек отношений предшествования (в конфигурации приложения на рисунке 1 такое пересечение имеет место в задаче τ_3).

Вместе с тем, не во всех приложениях, содержащих задачи с пересекающимися критическими интервалами, возможно возникновение цепочек отношений предшествования. Так, если в приложении на рисунке 1 задаче τ_3 назначить низший приоритет, то цепочка отношений предшествования не возникнет (будет достаточно использовать процедуру непосредственного наследования приоритетов). Необходимость использования транзитивного наследования приоритетов может возникнуть только для приложений, допускающих цепное блокирование задач. Метод проверки наличия в приложении условий для возникновения цепного блокирования представлен в [18].

8. Динамическое обнаружение и выход из ситуаций взаимного блокирования заданий. На структуру программных приложений, использующих традиционный вариант процедуры наследования приоритетов [16], накладываются два ограничения. Одно из них – отмеченное в разделе 3 требование вложенности пересекающихся критических интервалов. Другое структурное ограничение – требование гарантированной живучести. В ходе работы приложений, отвечающих требованию гарантированной живучести, невозможно возникновение ситуаций типа взаимного блокирования заданий – ситуаций, отличающихся тем, что в ориентированном графе, отражающем текущий состав отношений предшествования ${}_i\tau_x \succ {}_j\tau_y$ исполняемых заданий, возникают замкнутые маршруты (контуры). Классический пример структуры приложения из пяти задач с пятью ресурсами, не отвечающий требованию живучести, приведен в работе [19].

В системах реального времени использование приложений, не отвечающих требованию живучести, недопустимо. Отсюда следует, что программные комплексы для систем реального времени, ориентированные на использование традиционного варианта транзитивной процедуры наследования приоритетов, должны статически (на этапе проектирования) проверяться на живучесть. Объем вычислений для такой проверки растет экспоненциально с ростом числа взаимосвязанных задач. Для приложений, содержащих десятки взаимосвязанных задач, выполнение такой проверки может оказаться нереальным. В этом случае обеспечение живучести может обеспечиваться динамически путем встраивания процедуры транзитивного наследования приоритетов в механизм обработки исключительных ситуаций.

В ходе исполнения программного приложения, не отвечающего требованию живучести, возможно возникновение состояния, в котором процедура транзитивного наследования приоритетов обнаружит, задание, выявляемое на шаге 2, совпадает с тем заданием, которое обратилось к операции *unlock* (цепочка отношений предшествования

замыкается, что соответствует возникновению взаимного блокирования заданий). Для приложений, не отвечающих требованию живучести, шаг 2 процедуры транзитивного наследования должен быть дополнен проверкой на возникновение такого совпадения. Если при этом выполнение операции *unlock* встроено в механизм обработки исключительных ситуаций (*unlock* выполняется в рамках оператора *try*), то информация о попытке замыкания отношений предшествования может быть передана из контекста подсистемы управления заданиями в контекст приложения – в *catch*-ветвь оператора *try*. Таким образом приложение информируется о попытке взаимного блокирования заданий. Дальнейшие поведение приложения зависит от действий, заложенных проектировщиком приложения в *catch*-ветви. В частности, текущая задача может обойти ситуацию взаимного блокирования путем досрочного освобождения одного из уже занятых ею ресурсов. Тем самым обеспечивается динамическое обнаружение и выход из ситуаций взаимного блокирования заданий.

9. Заключение. В многозадачных программных приложениях реального времени требование целостности разделяемых задачами глобальных информационных ресурсов обеспечивается защитой критических интервалов по доступу к разделяемым ресурсам специальными синхронизирующими элементами типа мьютексов. При исполнении таких приложений возникают отношения предшествования между критическими интервалами взаимосвязанных задач. Возникающая в результате этого инверсия приоритетов может привести к нарушению предельных сроков выполнения задач. Предотвращение инверсии приоритетов обеспечивается либо непосредственной, либо транзитивной процедурой наследования приоритетов. Применение упрощенной непосредственной процедуры гарантирует предотвращение инверсии приоритетов в случае, если в структуре задач отсутствуют пересечения критических интервалов. При наличии таких пересечений могут возникать цепочки отношений предшествования. В этом случае предотвращение инверсии приоритетов достигается применением транзитивной процедуры. Предложенная в настоящей статье модификация транзитивной процедуры со встраиванием операции освобождения ресурса в механизм обработки исключительных ситуаций обеспечивает обнаружение и выход из ситуаций типа взаимного блокирования заданий.

Литература

1. *Давиденко К.Я.* Технология программирования АСУТП. Проектирование систем реального времени, параллельных и распределенных приложений // М.: Энергоатомиздат, 1985. 183 с.
2. *Baker T.* Multiprocessors EDF and Deadline Monotonic Schedulability Analysis // Proceedings of 24 IEEE Real-Time Systems Symposium. 2003. pp. 120–129.

3. *Таненбаум Э.* Современные операционные системы // СПб.: Питер. 2010. 1037 с.
4. *Сорокин С.В.* Системы реального времени: операционные системы // Современные технологии автоматизации. 1997. №2. С. 22–31.
5. *Liu C., Layland J.* Scheduling Algorithms for Multiprocessing in a Hard Real-Time Environment // Journal of the ACM. 1973. vol. 20. no. 1. pp. 46–61.
6. *Dhall S.K., Liu C.L.* On a Real-Time Scheduling Problem // Operating Research. 1978. vol. 26. no. 1. pp. 127–140.
7. *Никифоров В.В.* Выполнимость приложений реального времени на многоядерных процессорах // Труды СПИИРАН. 2009. Вып. 8. С. 255–284.
8. *Докучаев А.Н.* К оценке эффективности механизмов диспетчеризации мультипроцессорных систем реального времени с учетом влияния длительных блокировок // Программная инженерия. 2012. №9. С. 2–7.
9. *Anderson B.* Global Static-Priority Preemptive Multiprocessor Scheduling with Utilization Bound 38% // Proceedings of the7th International Conference on Principles of Distributed Systems. Egypt. 2008. pp. 73–88.
10. *Baker T.P., Cirinei M., Bertogna M.* EDZL scheduling analysis // Real-Time Systems 2008. vol. 40. no. 3. pp. 264–289.
11. *Baruah S.K.* Fairness in Periodic Real-Time Scheduling Algorithms // Proceedings of 16 IEEE Real-Time Symposium. 1995. pp. 200–209.
12. *Cho H., Ravindran B., Jensen D.* An Optimal Real-Time Scheduling Algorithm for Multiprocessors // Proceedings of the 27 IEEE Real-Time Symposium. 2006. pp. 101–110.
13. *Liu J.W.S.* Real-Time Systems // NJ: Prentice Hall. 2000. 590 p.
14. *Laplante P.A.* Real-Time Systems Design and Analysis // John Wiley & Sons, Inc. 2004. 530 p.
15. *Данилов М.В.* Методы планирования выполнения задач в системах реального времени // Программные продукты и системы. 2001. №4. С. 28–35.
16. *Sha L., Rajkumar R., Lehoczky J.P.* Priority Inheritance Protocols: An Approach to Real-Time Synchronization // IEEE Transactions on Computers. 1990. vol. 20. no. 9. pp.1175–1185.
17. *Никифоров В.В., Шкиртиль В.И.* Спецификация средствами языка XML системы интерфейсов в приложениях реального времени // Труды СПИИРАН. 2009. Вып 11. С.159–175.
18. *Никифоров В.В., Шкиртиль В.И.* Цепное блокирование взаимосвязанных задач в системах на многоядерных процессорах // Информационно-измерительные и управляющие системы. 2013. №9. С. 17–21.
19. *Dijkstra E.W.* Hierarchical ordering of sequential processes // Acta Informatica. 1971. vol. 1. no. 2. pp. 115–138.

References

1. Davidenko K.Ya. *Tekhnologiya programmirovaniya ASUTP. Proyektirovaniye sistem realnogo vremeni, parallelnyh i raspredelennyh prilozheniy* [Technology of programming with CAD/CAM. Design of real-time systems, parallel and distributed applications]. M. 1985. 183 p. (In Russ.).
2. Baker T. Multiprocessors EDF and Deadline Monotonic Schedulability Analysis. Proceedings of 24 IEEE Real-Time Systems Symposium. 2003. pp. 120–129.
3. Tannenbaum E. *Sovremennye operatsionnye sistemy* [Modern operating systems]. St.Petersburg. 2010. 1037 p. (In Russ.)
4. Sorokin S.V. [Real-time systems: operating systems]. *Sovremennye tekhnologii avtomatizatsii – Modern technologies of automation*. 1997. vol. 2. pp. 22–31. (In Russ.).
5. Liu C., Layland J. Scheduling Algorithms for Multiprocessing in a Hard Real-Time Environment. *Journal of the ACM*. 1973. vol. 20. no.1. pp. 46–61.

6. Dhall S.K., Liu C.L. On a Real-Time Scheduling Problem. *Operating Research*. 1978. vol. 26. no. 1. pp. 127–140.
7. Nikiforov V.V. [Feasibility of real-time applications on multi-core processors]. *Trudy SPIIRAN – SPIIRAS Proceedings*, issue 8. 2009. pp. 255–284. (In Russ.).
8. Dokuchayev A.N. [To esrimating the efficiency pf scheduling mechanisms of multiprocessor real-time systems taking into account long-term clinches.]. *Programmnaya inzheneriya – Software engineering*. 2012. vol. 9. pp. 2-–7. (In Russ.).
9. Anderson B. Global Static-Priority Preemptive Multiprocessor Scheduling with Utilization Bound 38%. Proceedings of the7th International Conference on Principles of Distributed Systems. Egypt. 2008. pp. 73–88.
10. Baker T.P., Cirinei M., Bertogna M. EDZL scheduling analysis. *Real-Time Systems*. 2008. vol. 40. no. 3. pp. 264–289.
11. Baruah S.K. Fairness in Periodic Real-Time Scheduling Algorithms. Proceedings of 16 IEEE Real-Time Symposium. 1995. pp. 200–209.
12. Cho H., Ravindran B., Jensen D. An Optimal Real-Time Scheduling Algorithm for Multiprocessors. Proceedings of the 27 IEEE Real-Time Symposium. 2006. pp. 101–110.
13. Liu J.W.S. Real-Time Systems. NJ: Prentice Hall. 2000. 590 p.
14. Laplante P.A. Real-Time Systems Design and Analysis. John Wiley & Sons, Inc. 2004. 530 p.
15. Danilov M.V. [Scheduling methods od task execution in real-time systems]. *Programmnye produkty i sistemy – Software products and systems*. 2001. vol. 4. pp. 28–35. (In Russ.).
16. Sha L., Rajkumar R., Lehoczky J.P. Priority Inheritance Protocols: An Approach to Real-Time Synchronization. *IEEE Transactions on Computers*. 1990. vol. 20. no. 9. pp. 1175–1185.
17. Nikiforov V.V, Shkirtil V.I. [Specifying the interface system in real-time applications with XML]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2009. vol. 11. pp.159–175. (In Russ.).
18. Nikiforov V.V, Shkirtil V.I. [Chained blocking of interrelated tasks in systems on multi-core processors]. *Informatsionno-izmeritelnye i upravlyushie sistemy – Informational-measuring and control systems*. 2013. vol. 9. pp. 17–21. (In Russ.).
19. Dijkstra E.W. Hierarchical ordering of sequential processes. *Acta Informatica*. 1971. vol. 1. no. 2. pp. 115–138.

Баранов Сергей Николаевич — д-р физ.-мат. наук, профессор, главный научный сотрудник лаборатории информационно-вычислительных систем и технологий программирования, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН), профессор, международная научная лаборатория Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики (ИТМО). Область научных интересов: технология программирования, формальные методы. Число научных публикаций — 100. snbaranov@googlemail.com; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7-812-328-0887.

Baranov Sergey Nikolaevich — Ph.D., Dr. Sci., professor, chief researcher of computing and information systems and programming technologies laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), professor, Saint Petersburg National Research University of Information Technologies, Mechanics and Optics (ITMO University). Research interests: software engineering, formal methods in software development. The number of publications — 100. snbaranov@googlemail.com; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7-812-328-0887.

Никифоров Виктор Викентьевич — д-р техн. наук, профессор, ведущий научный сотрудник лаборатории информационно-вычислительных систем и технологий программирования, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: системы реального времени, встроенные системы, операционные системы. Число научных публикаций — 110. nik@iias.spb.su; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7(812)3280887.

Nikiforov Victor Vikentievich — Ph.D., Dr. Sci., professor, leading researcher of computing and information systems and programming technologies laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: real-time systems, embedded systems, operating systems. The number of publications — 110. nik@iias.spb.su; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)3280887.

Поддержка исследований. Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01), университет ИТМО.

Acknowledgements. This work was partially financially supported by the Government of the Russian Federation, Grant 074-U01.

РЕФЕРАТ

Баранов С.Н., Никифоров В.В. **Транзитивное наследование приоритетов в многозадачных приложениях реального времени.**

Программные продукты с широким спектром функциональных возможностей для пользователя обычно строятся в виде приложений, состоящих из ряда задач, каждой из которых предоставляется доступ к различным активным (исполнительным) и пассивным (информационным) ресурсам. При создании многозадачных программных приложений возникают две проблемы: 1) определение порядка предоставления задачам ресурса процессора (процессорного времени), гарантирующего своевременность их исполнения; 2) обеспечение целостности глобальных информационных ресурсов (т.е., предотвращение одновременного доступа к нему двух и более заданий). Если первая проблема решается применением дисциплины планирования, обеспечивающей корректность работы данного приложения, то вторая решается через использование специальных синхронизационных механизмов, предотвращающих одновременных доступ задач к общим ресурсам. Требование целостности разделяемых задачами информационных ресурсов обеспечивается защитой критических интервалов по доступу к ним специальными синхронизирующими механизмами типа мьютексов. При этом возникают отношения предшествования между критическими интервалами взаимосвязанных задач. Возникающая в результате этого инверсия приоритетов может привести к нарушению предельных сроков выполнения задач. Предотвращение инверсии приоритетов обеспечивается либо непосредственной, либо транзитивной процедурой наследования приоритетов. Процедура непосредственного наследования гарантирует предотвращение инверсии приоритетов в случае, если в структуре задач отсутствуют пересечения критических интервалов. При наличии таких пересечений могут возникать цепочки отношений предшествования. В этом случае предотвращение инверсии приоритетов достигается применением процедуры транзитивного наследования. В случае попадания подмножества задач в состояние взаимного блокирования транзитивная процедура позволяет обнаруживать этот факт. Предложенная в настоящей статье модификация транзитивной процедуры со встраиванием операции освобождения ресурса в механизм обработки исключительных ситуаций обеспечивает обнаружение ситуаций типа взаимного блокирования заданий с возможностью соответствующей реакции на них.

SUMMARY

Baranov S.N., Nikiforov V.V. **Transitive Priority Inheritance in Real-Time Multi-Task Applications.**

Software products with a wide range of functionality are usually built as multitask applications which consist of a number of tasks with access to various active (computational) and passive (informational) resources. Two problems often arise when developing such multi-task applications: 1) how to define the order of allocating the processor resource (processor time) to tasks which ensures their on-time execution; 2) how to preserve the integrity of global informational resources (i.e., how to exclude simultaneous access of two or more tasks to such a resource). While the first problem is resolved through a scheduling mode, which ensures the correctness of application execution, the second one is resolved via special synchronization mechanisms which prevent simultaneous task access to shared resources. The integrity of global resources shared among tasks is provided by protecting critical intervals of shared resources access with special synchronization mechanisms of the mutex type, which gives rise to the relation of precedence among critical intervals of interrelated tasks. The resulting priority inversion may cause a violation of task deadlines. Such priority inversion may be prevented by either a direct or a transitive procedure of priority inheritance. The direct procedure prevents priority inversion only if there are no intersections of critical intervals in the task structure. When such intersections do occur, chains of precedence relations may arise. In this case, priority inversion may be prevented with the proposed advanced transitive procedure. In addition, this procedure allows detecting an exception of mutual task clinch which a subset of tasks is trapped in during execution, with an option for a pre-planned reaction on such exceptions.

УДК **004.896**

А.И. Мотиенко, С.М. Макеев, О.О. Басов

АНАЛИЗ И МОДЕЛИРОВАНИЕ ПРОЦЕССА ВЫБОРА ПОЛОЖЕНИЯ ДЛЯ ТРАНСПОРТИРОВКИ ПОСТРАДАВШЕГО НА ОСНОВЕ БАЙЕСОВСКИХ СЕТЕЙ ДОВЕРИЯ

Мотиенко А.И., Макеев С.М., Басов О.О. **Анализ и моделирование процесса выбора положения для транспортировки пострадавшего на основе байесовских сетей доверия.**

Аннотация. Целью любых аварийно-спасательных и других неотложных работ является спасение людей и оказание помощи пострадавшим, локализация аварий и устранение повреждений, препятствующих проведению спасательных работ, а также создание условий для последующего проведения восстановительных работ. При наличии факторов, угрожающих жизни и здоровью проводящих эти работы людей (спасателей, пожарных и др.) возникает объективная необходимость в применении автоматизированных робототехнических средств транспортировки пострадавших, а отсутствие соответствующего научно-методического и программно-алгоритмического инструментария обусловливает необходимость моделирования указанных средств. В работе представлена модель положения для транспортировки пострадавшего на основе байесовских сетей доверия.

Ключевые слова: робототехника, аварийно-спасательные роботы, человеко-машинное взаимодействие, транспортировка пострадавших, первая помощь, аварийно-спасательные работы, чрезвычайная ситуация, байесовские сети доверия.

Motienko A.I., Makeev S.M., Basov O.O. **Analysis and Modeling of the Process of a Choice of Position for Transportation of the Sufferer on the basis of Bayesian Belief Networks.**

Abstract. The purpose of any rescue and other emergency operations is to rescue of people and assistance to sufferer, localization of accidents and elimination of the damages interfering carrying out rescue efforts, and also creation of conditions for the subsequent carrying out recovery work. In the presence of the factors menacing to life and health of the people who are carrying out these works (rescuers, firefighters, etc.) there is an objective need for application of the automated robotic means of transportation of sufferer, and lack of the corresponding scientific and methodical and program and algorithmic tools causes need of modeling of the specified means. A model of position for transportation of the sufferer on the basis of Bayesian belief networks is presented in the paper.

Keywords: robotics, search-and-rescue, human-machine interaction, transportation of the sufferer, first aid, rescue work, emergency, Bayes belief network.

1. Введение. Согласно данным мировой статистики (начиная со второй половины XX века) первое место в перечне причин смерти трудоспособного населения в возрасте до 45 лет занимают травмы. В свою очередь, на втором месте (после дорожно-транспортных происшествий) среди причин такой смертности находятся стихийные бедствия и пожары. Среди стихийных бедствий в России наибольшее число людских потерь характерно для чрезвычайных ситуаций, связанных с наводнениями и землетрясениями. В пожарах в нашей стране ежегодно погибает 13-15 тысяч человек. Проведение аварийно-спасательных и других неотложных работ, направленных на

ликвидацию чрезвычайных ситуаций, является одной из основных задач Российской единой Системы предотвращения чрезвычайных ситуаций и Гражданской обороны [1].

Целью любых аварийно-спасательных и других неотложных работ является спасение людей и оказание помощи пострадавшим, локализация аварий и устранение повреждений, препятствующих проведению спасательных работ, а также создание условий для последующего проведения восстановительных работ. По данным Всемирной организации здравоохранения около 85 % среди погибших от травм умирают вследствие не оказанной, оказанной несвоевременно или неправильно первой помощи на месте чрезвычайной ситуации (и только 15 % погибают от несовместимых с жизнью повреждений) [2].

В соответствии с действующими нормативными документами при возникновении чрезвычайных ситуаций создаются временный штаб по ликвидации и предварительный план мероприятий. Последний включает себя:

a) предварительную разведку маршрутов движения формирований и участков предстоящих работ и уточнение ситуации в районе чрезвычайной ситуации;

b) дальнейшую наземную разведку, прокладку колонных путей и устройство проездов (проходов) в завалах и на заражённых участках, а также локализацию и тушение пожаров на путях движения формирований и участках работ;

c) локализацию аварий на коммунально-энергетических и технологических сетях;

d) розыск пострадавших и извлечение их из под завалов, повреждённых и горящих зданий, загазованных, задымлённых и затопленных помещений, санитарная обработка людей, обеззараживание их одежды, территории, сооружений, техники, воды и продовольствия;

e) оказание первой помощи пострадавшим и транспортировка их в лечебные учреждения.

Аварийно-спасательные работы характеризуются наличием факторов, угрожающих жизни и здоровью проводящих эти работы людей (спасателей, пожарных и др.), и требуют специальной подготовки, экипировки и оснащения. Свести к минимуму степень риска для спасателей позволяет использование так называемых безлюдных технологий – автоматизированных робототехнических комплексов и средств [3].

Рынок робототехники в России существует и развивается уже более 10 лет. Современные разработки применяются в разных

областях: от социально-бытовой до военно-технической, как в штатных ситуациях, так и в экстремальных. Робототехническое оборудование используется при проведении аварийно-спасательных работ, в медицине, в ходе боевых действий и антитеррористических операций, разведки, охраны, разминирования и пр., обеспечивая высокую эффективность проводимых работ и максимальную безопасность здоровью и жизни человека.

Последняя задача (по п. е)) во многом сложнее предыдущих. Автоматизированные робототехнические средства должны обеспечить выполнение мероприятий по оказанию первой помощи [4] и непосредственно транспортировку пострадавшего. Однако, для определения признаков жизни (наличия сознания) и осмотра пострадавшего требуется наличие у робототехнических средств соответствующих датчиков (температуры, давления, влажности) и систем (компьютерного зрения, анализа и синтеза речи для опроса пострадавшего и др.), номенклатура и конфигурация которых должна определяться с учетом массогабаритных и энергетических характеристик робота. Для определения оптимального положения тела пострадавшего необходим соответствующий научно-методический инструментарий, позволяющий определить его в условиях неполных данных относительно признаков травм, состояния пострадавшего и симптомов заболевания.

Таким образом, в описанной предметной области, с одной стороны, наблюдается объективная необходимость в применении автоматизированных робототехнических средств, с другой – отсутствие соответствующего научно-методического аппарата и программно-алгоритмических решений для обеспечения возможности транспортировки пострадавших. Поэтому целью данного исследования является анализ и моделирование процесса выбора положения для транспортировки.

2. Обзор методов формализации процесса выбора положения для транспортировки пострадавшего. Указанный процесс является трудноформализуемым, при разработке модели могут быть использованы: деревья решений, метод анализа иерархий, продукционная модель представления знаний, нечеткие когнитивные модели, нейронные сети, байесовский вывод, а также байесовские сети доверия (БСД).

Недостатком *метода деревьев решений* является то, что они хорошо подходят для задач с небольшим числом возможных результатов, но неприменимы к наборам данных, где число возможных исходов достаточно велико [5]. *Метод анализа иерархий* основан на критерии отбора экспертов, поэтому нуждается в использовании дополнительных процедур с соответствующими алгоритмами,

требующими специального изучения [6]. Существенным недостатком *продукционной модели* является то, что при накоплении достаточно большого числа (порядка нескольких сотен) продукций они начинают противоречить друг другу. Указанный недостаток делает невозможным применение продукционной модели в процессе функционирования сложной слабоформализуемой системы [7]. К недостаткам *когнитивного моделирования* следует отнести невозможность доказательства адекватности разработанной модели [8]. Значительную сложность при применении *искусственных нейронных сетей* представляют выбор ее архитектуры для решаемой задачи и низкая скорость процесса обучения [9].

Относительно свободным от указанных недостатков при решении задачи моделирования процесса выбора положения для транспортировки пострадавшего является байесовский подход, основанный на принципе максимального использования имеющейся априорной информации, ее непрерывного пересмотра и переоценки с учетом получаемых выборочных данных об исследуемом явлении или процессе. Так как байесовский метод основан на наблюдениях, то с его помощью можно последовательно вычислить вероятность истинной гипотезы. При этом новые наблюдения или решения применяются для модификации априорных вероятностей, которые в свою очередь необходимы для вычисления апостериорных вероятностей гипотез. Кроме того, в отсутствие эмпирических данных рассматриваемый метод обеспечивает использование субъективных вероятностных оценок для априорных гипотез [10].

Байесовский вывод может быть развит на сети (графы), в которых узлы (вершины) представлены случайными переменными (наблюдениями или состояниями) различных типов, а связи между ними (ребра) показывают их вероятностные зависимости. Преимуществами байесовских сетей доверия (БСД) над остальными методами являются способность моделировать сложные зависимости между узлами и возможность простого интегрирования изменяющейся во времени разнородной информации [11].

БСД используются для рассуждений в условиях неопределенности и все чаще применяются в диагностике заболеваний, выборе оптимального курса лечения пациента, предсказании исхода заболевания, построении моделей заболеваний в клинической эпидемиологии [12]. Данный научно-методический аппарат позволяет комбинировать имеющиеся статистические данные о характеристиках здоровья пациентов в дополнение к экспертной информации, которую предоставляют врачи-специалисты. Кроме того, БСД (по сравнению с другими методами) позволяют моделировать

возможность возникновения нескольких заболеваний, а ее элементы имеют достаточно простую интерпретацию [13]. Важным достоинством моделей, опирающихся на принципы искусственного интеллекта, является возможность автоматического обучения структуры модели, то есть даже если первоначальная структура модели была неполной, то имеется возможность улучшить модель при помощи поступающих данных [14].

БСД обычно представляются графически в виде направленного ациклического графа и таблиц условных вероятностей для узлов графа, соответствующих определенным переменным [15]. Процесс работы с ними заключается в выполнении двух основных операций: обучения (формирования таблиц условных вероятностей) БСД на основе имеющихся данных о переменных сети [16, 17] и непосредственного использования БСД для вычисления различных вероятностей, связанных с переменными сети.

3. Формирование структуры и обучение байесовской сети доверия для моделирования процесса выбора положения для транспортировки пострадавшего. Структура разработанной БСД представлена признаками травм, непосредственно травмами, соответствующими им положениями для транспортировки пострадавшего и взаимосвязями между ними (рисунок 1). При этом узлы графа БСД введены, исходя из [2], на основе признаков травм, которые можно определить визуально, путем опроса или при помощи несложных манипуляций. В предложенной модели можно выделить следующие подграфы (для наглядности представления повторяющиеся вершины графа обозначены пунктиром):

– подграф G_1 «Положение для транспортировки», определяющий взаимосвязи между положениями для транспортировки (таблица 1) и соответствующими им травмами (таблица 2);

– подграф G_2 «Повреждение позвоночника» с таблицей условных вероятностей $p\left(x_6 \left| \tilde{x}_{24} \tilde{x}_{25} \tilde{x}_{26} \tilde{x}_{27} \tilde{x}_{28} \tilde{x}_{29} \tilde{x}_{30} \right.\right)$;

Таблица 1. Оптимальные позы для транспортировки пострадавших

Показатель	Наименование
x_1	На спине
x_2	На животе (с валиком под грудью и головой)
x_3	На правом боку
x_4	Сидя (с поднятой вверх рукой при ампутации)
x_5	Полусидячее положение со склоненной на грудь головой

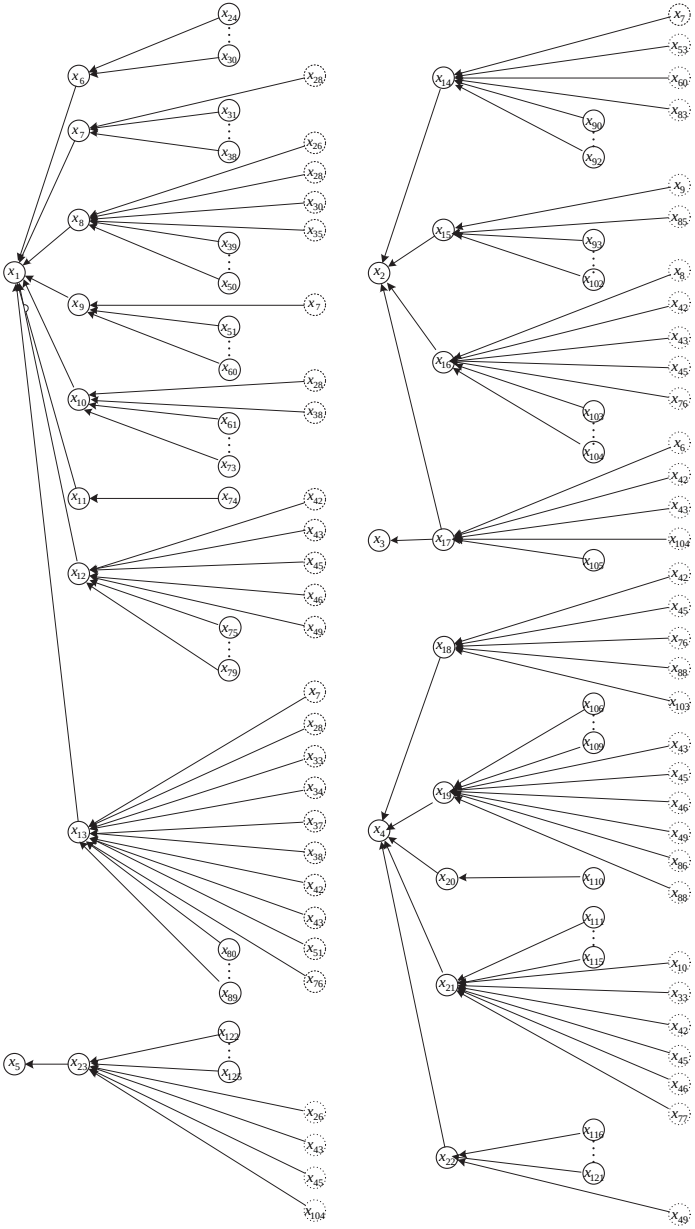


Рис. 1. Модель положения для транспортировки пострадавшего на основе БСД

Таблица 2. Перечень наиболее распространенных травм

Показатель	Наименование
x_6	Повреждение позвоночника
x_7	Шоковое состояние
x_8	Переломы костей таза и нижних конечностей
x_9	Сотрясение головного мозга
x_{10}	Травмы груди
x_{11}	Ампутация нижних конечностей
x_{12}	Травмы передней части головы и лица
x_{13}	Травмы органов брюшной полости
x_{14}	Кровопотеря
x_{15}	Травмы затылочной части головы
x_{16}	Травмы ягодиц, тыльной поверхности ног
x_{17}	Травмы спины
x_{18}	Ушибы, порезы, ссадины
x_{19}	Травмы плечевого пояса
x_{20}	Ампутированная верхняя конечность,
x_{21}	Травмы глаза, груди, дыхательных путей
x_{22}	Травмы верхних конечностей
x_{23}	Травмы шеи

– подграф G_3 «Шоковое состояние» с таблицей условных вероятностей $p\left(x_7 \left| \tilde{x}_{28}\tilde{x}_{31}\tilde{x}_{32}\tilde{x}_{33}\tilde{x}_{34}\tilde{x}_{35}\tilde{x}_{36}\tilde{x}_{37}\tilde{x}_{38}\right.\right)$;

– подграф G_4 «Переломы костей таза и нижних конечностей» с таблицей условных вероятностей $p\left(x_8 \left| \tilde{x}_{26}\tilde{x}_{28}\tilde{x}_{35}\tilde{x}_{39}\tilde{x}_{40}\tilde{x}_{41}\tilde{x}_{42}\tilde{x}_{43}\tilde{x}_{44}\tilde{x}_{45}\tilde{x}_{46}\tilde{x}_{47}\tilde{x}_{48}\tilde{x}_{49}\tilde{x}_{50}\right.\right)$;

– подграф G_5 «Сотрясение головного мозга» с таблицей условных вероятностей $p\left(x_9 \left| \tilde{x}_7\tilde{x}_{51}\tilde{x}_{52}\tilde{x}_{53}\tilde{x}_{54}\tilde{x}_{55}\tilde{x}_{56}\tilde{x}_{57}\tilde{x}_{58}\tilde{x}_{59}\tilde{x}_{60}\right.\right)$;

– подграф G_6 «Травмы груди» с таблицей условных вероятностей $p\left(x_{10} \left| \tilde{x}_{28}\tilde{x}_{61}\tilde{x}_{62}\tilde{x}_{63}\tilde{x}_{64}\tilde{x}_{65}\tilde{x}_{66}\tilde{x}_{67}\tilde{x}_{68}\tilde{x}_{69}\tilde{x}_{70}\tilde{x}_{71}\tilde{x}_{72}\tilde{x}_{73}\tilde{x}_{74}\right.\right)$;

– подграф G_7 «Ампутированная нижняя конечность» с таблицей условных вероятностей $p\left(x_{11} \left| \tilde{x}_{75}\right.\right)$;

– подграф G_8 «Травмы передней части головы и лица» с таблицей условных вероятностей $p\left(x_{12} \left| \tilde{x}_{42}\tilde{x}_{43}\tilde{x}_{45}\tilde{x}_{46}\tilde{x}_{76}\tilde{x}_{77}\tilde{x}_{78}\tilde{x}_{79}\tilde{x}_{80}\tilde{x}_{81}\right.\right)$;

– подграф G_9 «Травмы органов брюшной полости» с таблицей условных вероятностей $p\left(x_{13} \left| \tilde{x}_7\tilde{x}_{28}\tilde{x}_{33}\tilde{x}_{34}\tilde{x}_{37}\tilde{x}_{38}\tilde{x}_{42}\tilde{x}_{43}\tilde{x}_{78}\tilde{x}_{83}\tilde{x}_{84}\tilde{x}_{85}\tilde{x}_{86}\tilde{x}_{87}\tilde{x}_{89}\tilde{x}_{90}\tilde{x}_{91}\tilde{x}_{92}\right.\right)$;

– подграф G_{10} «Кровопотеря» с таблицей условных вероятностей $p\left(x_{14}\left|\tilde{x}_7\tilde{x}_{53}\tilde{x}_{60}\tilde{x}_{85}\tilde{x}_{93}\tilde{x}_{94}\tilde{x}_{95}\right.\right)$;

– подграф G_{11} «Травмы затылочной части головы» с таблицей условных вероятностей $p\left(x_{15}\left|\tilde{x}_9\tilde{x}_{87}\tilde{x}_{96}\tilde{x}_{97}\tilde{x}_{98}\tilde{x}_{99}\tilde{x}_{100}\tilde{x}_{101}\tilde{x}_{102}\tilde{x}_{103}\tilde{x}_{104}\tilde{x}_{105}\right.\right)$;

– подграф G_{12} «Травмы ягодиц, тыльной поверхности ног» с таблицей условных вероятностей $p\left(x_{16}\left|\tilde{x}_8\tilde{x}_{42}\tilde{x}_{43}\tilde{x}_{45}\tilde{x}_{78}\tilde{x}_{106}\tilde{x}_{107}\right.\right)$;

– подграф G_{13} «Травмы спины» с таблицей условных вероятностей $p\left(x_{17}\left|\tilde{x}_6\tilde{x}_{42}\tilde{x}_{43}\tilde{x}_{107}\right.\right)$;

– подграф G_{14} «Ушибы, порезы, ссадины» с таблицей условных вероятностей $p\left(x_{18}\left|\tilde{x}_{42}\tilde{x}_{45}\tilde{x}_{78}\tilde{x}_{106}\tilde{x}_{109}\right.\right)$;

– подграф G_{15} «Травмы плечевого пояса» с таблицей условных вероятностей $p\left(x_{19}\left|\tilde{x}_{43}\tilde{x}_{45}\tilde{x}_{46}\tilde{x}_{76}\tilde{x}_{89}\tilde{x}_{110}\tilde{x}_{111}\tilde{x}_{112}\tilde{x}_{113}\right.\right)$;

– подграф G_{16} «Ампутированная верхняя конечность» с таблицей условных вероятностей $p\left(x_{20}\left|\tilde{x}_{114}\right.\right)$;

– подграф G_{17} «Травмы глаза, груди, дыхательных путей» с таблицей условных вероятностей $p\left(x_{21}\left|\tilde{x}_{10}\tilde{x}_{33}\tilde{x}_{42}\tilde{x}_{45}\tilde{x}_{46}\tilde{x}_{115}\tilde{x}_{116}\tilde{x}_{117}\tilde{x}_{118}\tilde{x}_{119}\tilde{x}_{120}\right.\right)$;

– подграф G_{18} «Травмы верхних конечностей» с таблицей условных вероятностей $p\left(x_{22}\left|\tilde{x}_{76}\tilde{x}_{121}\tilde{x}_{122}\tilde{x}_{123}\tilde{x}_{124}\tilde{x}_{125}\tilde{x}_{126}\right.\right)$;

– подграф G_{19} «Травмы шеи» с таблицей условных вероятностей $p\left(x_{23}\left|\tilde{x}_{43}\tilde{x}_{45}\tilde{x}_{54}\tilde{x}_{107}\tilde{x}_{127}\tilde{x}_{128}\tilde{x}_{129}\tilde{x}_{130}\right.\right)$.

Обозначение $\tilde{x}$ используется для указания, что на этом месте в формуле может стоять как сама пропозициональная формула x, так и ее отрицание $\overline{x}$ [18].

Признаки травм, используемые для задания узлов разработанной БСД, и способы их определения (путем осмотра, опроса, манипуляций) представлены в таблице 3.

Для подграфа G_1 таблица условных вероятностей имеет вид:

$$p\left(x_1\left|\tilde{x}_6\tilde{x}_7\tilde{x}_8\tilde{x}_9\tilde{x}_{10}\tilde{x}_{11}\tilde{x}_{12}\tilde{x}_{13}\right.\right)=1,$$

если имеет место хотя бы одна пропозициональная формула x_i, $i=6...13$;

Таблица 3. Признаки травм и способы их определения

Показатель	Наименование	Осмотр	Опрос	Манипуляции
x_{24}	снижение чувствительности, жжение		+	
x_{25}	неестественное положение шеи и спины	+		
x_{26}	локализация боли		+	
x_{27}	нарушение двигательной функции	+	+	
x_{28}	пониженное АД			+
x_{29}	онемение		+	
x_{30}	нарушение функций тазовых органов	+		
x_{31}	потеря сознания	+		
x_{32}	возбуждение, сменяющееся заторможенностью	+		
x_{33}	учащенное дыхание			+
x_{34}	учащенный пульс			+
x_{35}	тахикардия			+
x_{36}	беспамятство		+	
x_{37}	потливость			+
x_{38}	бледность	+		
x_{39}	деформация тазовой области, конечностей	+		
x_{40}	визуальное укорочение конечности	+		
x_{41}	нарушение подвижности нижних конечностей	+	+	
x_{42}	гематомы	+		
x_{43}	раны	+		
x_{44}	уменьшение движений в тазобедренном суставе, конечностях	+	+	
x_{45}	кровотечение из раневой поверхности	+		
x_{46}	отек мягких тканей в области травмы	+		
x_{47}	ротация конечности	+		
x_{48}	патологическая подвижность	+		
x_{49}	костные отломки	+		
x_{50}	вынужденное положение конечности	+		
x_{51}	заметные повреждения	+		
x_{52}	кровь из носа	+		
x_{53}	головокружение		+	
x_{54}	тошнота		+	
x_{55}	слабость		+	
x_{56}	свето- и звукобоязнь		+	
x_{57}	нарушение координации движений	+		
x_{58}	головная боль		+	

Продолжение таблицы 3.

Показатель	Наименование	Осмотр	Опрос	Манипуляции
x_{59}	расширенные/суженные зрачки	+		
x_{60}	шум в ушах		+	
x_{61}	необычные дыхательные шумы			+
x_{62}	парадоксальное дыхание			+
x_{63}	сосущие раны грудной клетки	+		
x_{64}	кардиалгия		+	
x_{65}	набухшие не пульсирующие шейные вены	+		
x_{66}	болевой синдром, усиливающийся при кашле		+	
x_{67}	подкожная эмфизема	+		
x_{68}	торакалгия		+	
x_{69}	цианоз	+		
x_{70}	одностороннее дыхание			+
x_{71}	дыхательные движения короткие и поверхностные	+		
x_{72}	травматическая асфиксия	+		
x_{73}	симптом "декольте"	+		
x_{74}	отсутствие нижней конечности	+		
x_{75}	асимметрия лица	+		
x_{76}	ссадины	+		
x_{77}	нарушение зрения		+	
x_{78}	деформация передней части головы и лица	+		
x_{79}	целостность глаз, носа	+		
x_{81}	напряжение мышц брюшной стенки	+		
x_{82}	ограничение дыхательных движений брюшной стенки	+		
x_{83}	жажда		+	
x_{84}	повышение температуры			+
x_{85}	рвота	+		
x_{86}	кровоизлияния	+		
x_{87}	раневые поверхности	+		
x_{88}	локальная припухлость и болезненность	+	+	
x_{89}	отечность в области промежности	+		
x_{90}	сонливость (зевота)		+	
x_{91}	брадикардия			+
x_{92}	круги перед глазами		+	
x_{93}	психомоторное возбуждение	+	+	
x_{94}	односторонний мидриаз	+		

Окончание таблицы 3.

Показатель	Наименование	Осмотр	Опрос	Манипуляции
x_{95}	снижение реакции зрачков на свет	+		
x_{96}	кома	+		
x_{97}	гемипарез		+	
x_{98}	повышение артериального давления			+
x_{99}	фокальные эпилептические припадки	+		
x_{101}	непроизвольные колебательные движения глаз	+		
x_{102}	ограничение взора вверх	+		
x_{103}	порезы	+		
x_{104}	ушибы	+		
x_{106}	укорочение предплечья	+		
x_{107}	опущение и смещение кпереди плеча	+		
x_{108}	ограничение, болезненность движений		+	
x_{109}	деформация	+		
x_{100}	нарушение речи		+	
x_{110}	отсутствие верхней конечности	+		
x_{111}	нарушение целостности области глаза	+		
x_{112}	наличие инородного тела в области глаза	+		
x_{113}	уряженное дыхание			+
x_{114}	одышка			+
x_{115}	боли при дыхании		+	
x_{116}	хруст			+
x_{117}	травмированная часть конечностей изменена	+		
x_{118}	неправильное положение верхней конечности	+		
x_{119}	боли в неподвижном состоянии		+	
x_{120}	свобода движений только в месте, где нет суставов	+		
x_{121}	ограниченность в движениях	+		
x_{122}	невозможность поворачивать голову	+		
x_{123}	выраженный кифоз (выпуклый кзади)	+		
x_{124}	вынужденное положение шеи	+		
x_{125}	фиксированное положение головы	+		

$$p\left(x_2 \left| \tilde{x}_{14}\tilde{x}_{15}\tilde{x}_{16}\overline{x}_{17}\right.\right) = 1,$$

если имеет место хотя бы одна пропозициональная формула x_i, $i = 14...16$;

$p\left(x_2 \left| x_{14} x_{15} x_{16} x_{17}\right.\right) = 1; \quad p\left(x_2 \left| \overline{x}_{14} \overline{x}_{15} \overline{x}_{16} x_{17}\right.\right) = 0,5; \quad p\left(x_2 \left| \tilde{x}_{14} \tilde{x}_{15} \tilde{x}_{16} x_{17}\right.\right) = 0,75,$

при любом другом сочетании пропозициональных формул x_i и их отрицаний $\overline{x}_i$ $(i = 14...16)$;

$$p\left(x_3 \left| x_{17}\right.\right) = 1; \; p\left(x_3 \left| \overline{x}_{17}\right.\right) = 0;$$
$$p\left(x_4 \left| \tilde{x}_{18} \tilde{x}_{19} \tilde{x}_{20} \tilde{x}_{21} \tilde{x}_{22}\right.\right) = 1,$$

если имеет место хотя бы одна пропозициональная формула x_i, $i = 18...22$;

$$p\left(x_5 \left| x_{23}\right.\right) = 1; \; p\left(x_5 \left| \overline{x}_{23}\right.\right) = 0$$

Для подграфа G_5 на начальном этапе формирования таблицы условных вероятностей допускается независимость появления событий $\left(x_7, x_{51}...x_{60}\right)$ и равнозначность их вклада в формирование апостериорной вероятности события x_9 «Сотрясение головного мозга». Тогда:

$$p\left(x_9 \left| \tilde{x}_7 \tilde{x}_{51} \tilde{x}_{53} \tilde{x}_{54} \tilde{x}_{55} \tilde{x}_{56} \tilde{x}_{57} \tilde{x}_{58} \; \tilde{x}_{59} \tilde{x}_{60}\right.\right) = 0,1, \qquad (1)$$

если в (1) имеет место одна пропозициональная формула x_i, $i = 7, 51, 53, 54...60$;

$$p\left(x_9 \left| \tilde{x}_7 \tilde{x}_{51} \tilde{x}_{53} \tilde{x}_{54} \tilde{x}_{55} \tilde{x}_{56} \tilde{x}_{57} \tilde{x}_{58} \; \tilde{x}_{59} \tilde{x}_{60}\right.\right) = 0,2, \qquad (2)$$

если в (2) имеет место две пропозициональных формулы x_i, $i = 7, 51, 53, 54...60$; и так далее, вплоть до случая:

$$p\left(x_9 \left| x_7 x_{51} x_{53} x_{54} x_{55} x_{56} x_{57} x_{58} \; x_{59} x_{60}\right.\right) = 1.$$

Аналогичный подход справедлив для всех подграфов $G_2,...,G_{19}$ травм $\left(x_6...x_{23}\right)$. Уточнение таблиц условных вероятностей для них осуществляется на основе экспертной информации, имеющихся медицинских исследований, направленных на выявление подобных взаимосвязей между элементами процесса диагностики травм, накопленных статистических данных.

4. Анализ процесса выбора положения для транспортировки пострадавшего. Процедура вычисления вероятностей исходов одной

или нескольких переменных БСД на основе таблиц условных вероятностей и известных данных (свидетельств) о значениях признаков травм называется опросом сети. Данная процедура выполняется как на этапе обучения сети, так и при ее непосредственном использовании, поэтому эффективность работы БСД во многом зависит от используемого алгоритма опроса и его реализации. Все известные методы опроса БСД можно условно разделить на две категории: алгоритмы, использующие представление БСД в более удобной для опроса форме (называемые также алгоритмами кластеризации), и приближенные алгоритмы, оперирующие стохастическими методами вычислений [19–21].

В проведенном исследовании применен алгоритм опроса БСД, использующий представление (кластеризацию) исходной сети в виде так называемого дерева сочленений (junction tree). Такой подход позволяет перейти от опроса сети общего вида к работе с древовидным графом, что существенно сокращает время вычислений, избавляя от необходимости во многих промежуточных расчетах. Данный алгоритм имеет ряд преимуществ:

1) использование дерева сочленений возможно для сетей любой топологической сложности, что делает этот алгоритм универсальным и применимым к очень широкому кругу задач;

2) в отличие от стохастических алгоритмов опроса дерево сочленений позволяет получить точные, а не приближенные значения требуемых вероятностей, при этом алгоритм обладает достаточно высокой скоростью работы.

Во всех популярных программах работы с БСД именно алгоритм дерева сочленений (в той или иной реализации) является основным алгоритмом опроса. Кроме того, для ряда задач, предполагающих работу с БСД и требующих точных результатов при сложной топологии сети, не существует приемлемого пути опроса, не использующего представления сети в виде дерева сочленений [22]. С учетом широкой распространенности и очевидной перспективности использования данного алгоритма, в настоящем исследовании проведено моделирование процесса выбора положения для транспортировки пострадавшего с использованием программы «Netica» [23].

На рисунке 2 приведен пример ее использования для вычисления вероятности для показателя x_1 (положение для транспортировки – на спине). Наиболее вероятное положение для транспортировки (рисунок 3) определялось как значение множества допустимых положений, доставляющее максимум вероятности

наличия травмы (травм), при условии конкретного набора свидетельств $(x_{24}...x_{130})$, включающего в себя признаки травм, симптомы и другие показатели (таблица 3).

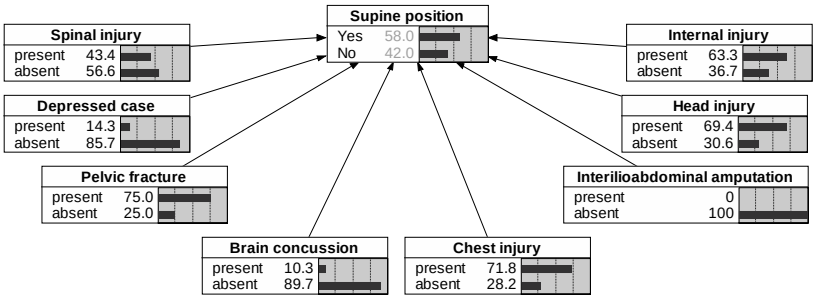


Рис. 2. Моделирование апостериорного вывода на подграфе G_1 «Положение для транспортировки» в программе «Netica»

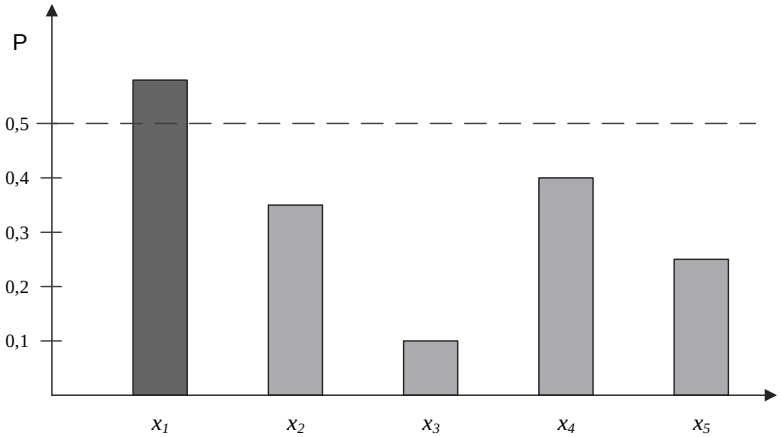


Рис. 3. Определение наиболее вероятного положения для транспортировки пострадавшего

5. Выводы. Результаты моделирования свидетельствуют, что разработанная БСД обеспечивает возможность решения задач вероятностного прогнозирования, базируясь на субъективных и неполных данных. Первые из них формируются в результате опроса пострадавшего, вторые – на основе систем компьютерного зрения (осмотр) и датчиков различного назначения (манипуляции), устанавливаемых на специализированных роботах.

Вероятностные оценки наблюдаемых свойств (свидетельств) дают фактические величины, позволяющие получить возможность пересмотра доверия к оценкам ненаблюдаемых свойств (травм и положений для транспортировки). Модель выбора положения для транспортировки пострадавшего дает возможность:

– установить, какие свидетельства являются наиболее важными по степени влияния на решение, составить их список и ранжировать по степени важности;

– выявить переменные x_i $(i = 24...125)$, не дающие оснований для выводов;

– создать упорядоченный список шагов, наиболее эффективно приводящих к ясному решению относительно положения для транспортировки (например, список вопросов, которые должны быть заданы пострадавшему при анализе ситуации, перечень и последовательность применения датчиков и т.п.).

Исследование указанных возможностей позволит модифицировать разработанную модель выбора положения для транспортировки пострадавшего, синтезировать алгоритмы оценки признаков травм, состояния человека и оптимизировать алгоритм вероятностного вывода.

Разрабатываемые средства ориентированы на развитие спасательной робототехники на основе интеллектуального программно-аппаратного обеспечения человеко-машинного взаимодействия [24–27].

Литература

1. *Мотиенко А.И., Басов О.О.* Применение автоматизированных робототехнических средств транспортировки для оказания первой помощи пострадавшим // Сборник научных статей 2-й Международной молодежной научно-технической конференции «Прогрессивные технологии и процессы». Курск. 2015. Т. 2. С. 216–220.
2. *Коннова Л.А., Балабанов В.А., Артамонова Г.К.* Основы первой помощи: учебник для курсантов, студентов и слушателей высших учебных заведений, обучающихся по направлению подготовки бакалавров «Техносферная безопасность» и специальности «Пожарная безопасность» / Под общей ред. О. М. Латышева // СПб.: Санкт-Петербургский университет ГПС МЧС России. 2015. 162 с.
3. *Мотиенко А.И., Ронжин А.Л., Павлюк Н.А.* Современные разработки аварийно-спасательных роботов: возможности и принципы их применения // Научный вестник НГТУ. 2015. № 3(60). С. 147–165.
4. Приказ Минздравсоцразвития РФ № 477н «Об утверждении перечня состояний, при которых оказывается первая помощь, и перечня мероприятий по оказанию первой помощи». 2012. 3 с.

5. *Михеев М.Ю., Котякова В.А., Володина Е.А., Баннов В.Я.* Применение «дерева решений» для анализа состояния сложных систем // Труды международного симпозиума Надёжность и качество. Пенза: ПГУ. 2012. Т. 2. С. 401–403.
6. *Тутыгин А.Г., Коробов В.Б.* Преимущества и недостатки метода анализа иерархий // Известия РГПУ им. А. И. Герцена. 2010. Т. 1. С. 108–115.
7. Продукционная модель знаний. URL: http://www.aiportal.ru/articles/knowledge-models/production-model.html (дата обращения: 29.10.2015).
8. *Кузькин А.А.* Методика обеспечения устойчивости стратегии развития информационных технологий на предприятии в условиях неопределенности воздействия среды // Дисс. к.т.н. Спб. 2015. 116 с.
9. *Басов О.О., Карпов А.А., Саитов И.А.* Методологические основы синтеза полимодальных инфокоммуникационных систем государственного управления: монография // Орёл: Академия ФСО России. 2015. 271 с.
10. *Мусина В.Ф.* Байесовские сети доверия как вероятностная графическая модель для оценки медицинских рисков // Труды СПИИРАН. 2013. Вып. 24. С. 135–151.
11. *Pitsikalis V., Katsamanis A., Papandreou G., Maragos P.* Adaptive multimodal fusion by uncertainty compensation // In Proceedings of the Ninth International Conference on Spoken Language Processing. Pittsburgh. 2006.
12. *Wasyluk H., Onisko A., Druzdzel M.J.* Support of diagnosis of liver disorders based on a causal Bayesian network model // Medical Science Monitor. 2001. vol. 7. pp. 327–332.
13. *Lacave C., Diez F.J.* Knowledge Acquisition in PROSTANET – A Bayesian network for diagnosis prostate cancer // Knowledge-Based Intelligent Information and Engineering Systems. 2003. LNCS 2774. pp. 1345–1350.
14. *Wiegerincka W.A.J.J., et al.* Approximate inference for medical diagnosis // Pattern Recognition Letters. 1999. vol. 20. no. 11–13. pp. 1231–1239.
15. *Jensen F.V., Nielsen T.D.* Bayesian networks and decision graphs // New York: Springer. 2007.
16. *Dempster A., Laird N., Rubin D.* Maximum likelihood from incomplete data via the EM algorithm // J. of the Royal Statistical Society. 1997. vol. 39. no. 1. pp. 1–38.
17. *Bender J., Koller D., Russel R., Kanazava K.* Adaptive probabilistic networks with hidden variables // Machine Learning. 1997. vol. 29. no. 2–3. pp. 213–244.
18. *Тулупьев А.Л., Сироткин А.В., Николенко С.И.* Байесовские сети доверия. Логико-вероятностный вывод в ациклических направленных графах // СПб.: Изд-во СПбГУ. 2009. 400 с.
19. *Henrion M.* Propagating uncertainty in Bayesian networks by logic sampling // Uncertainty in Artificial Intelligence. Amsterdam 1988. vol 2. pp. 149–163.
20. *Fung R., Chang K.-C.* Weighting and integrating evidence for stochastic simulation in Bayesian networks // Proc. ofthe Fifth Conference on Uncertainty in Artificial Intelligence (UAI-89). Amsterdam. 1989. pp. 475–482.
21. *Shachter R., Peot M.* Simulation approaches to general probabilistic inference on belief networks // Proc. of the Fifth Workshop on Uncertainty in Artificial Intelligence (UAI-89). Amsterdam. 1989. pp. 311–318.
22. *Масленников Е.Д., Сулимов В.Б.* Предсказания на основе байесовских сетей доверия: алгоритм и программная реализация // Вычислительные методы и программирование. 2010. Т. 11. № 2. С. 222–235.
23. Netica. URL: http://www.norsys.com/ (дата обращения: 12.09.2015).
24. *Ронжин А.Л., Юсупов Р.М.* Многомодальные интерфейсы автономных мобильных робототехнических комплексов // Известия ЮФУ. Технические науки. 2015. № 1(162). С. 195–206.
25. *Басов О.О.* Принципы построения полимодальных инфокоммуникационных систем на основе многомодальных архитектур абонентских терминалов // Труды СПИИРАН. 2015. Вып. 39. С. 109–122.

26. *Козыренко Н.С., Мещеряков Р.В., Ходашинский И.А., Ануфриева Н.Ю.* Математическое и алгоритмическое обеспечение оценки состояния здоровья человека // Труды СПИИРАН. 2014. Вып. 33. С. 117–146.
27. *Карпов А.А., Ронжин А.Л.* Многомодальные интерфейсы в автоматизированных системах управления // Известия высших учебных заведений. Приборостроение. 2005. Т. 48. № 7. С. 9-14.

References

1. Motienko A.I., Basov O.O. [Application of automated robotic means of transportation for first aid to sufferer]. *Sbornik nauchnyh statej 2-j Mezhdunarodnoj molodezhnoj nauchno-tehnicheskoj konferencii "Progressivnye tehnologii i process"* [Collected papers of International youth scientific and technical conference "Progressive Technologies and Processes"] Kursk. 2015. vol. 2. pp. 216–220. (In Russ.).
2. Konnova L.A., Balabanov V.A., Artamonova G.K. *Osnovy pervoj pomoshhi: uchebnik dlja kursantov, studentov i slushatelej vysshih uchebnyh zavedenij, obuchajushhihsja po napravleniju podgotovki bakalavrov «Tehnosfernaja bezopasnost'» i special'nosti «Pozharnaja bezopasnost'». Pod obshhej red. O.M. Latysheva* [Basics of First Aid: a textbook for students and students of higher educational institutions enrolled in the bachelor degree program "Technosphere safety" and the specialty "Fire safety". Edited by O.M. Latyshev]. SPb.: Sankt-Peterburgskij universitet GPS MChS Rossii. 2015. 162 p. (In Russ.).
3. Motienko A.I., Ronzhin A.L., Pavljuk N.A. [The modern development of rescue robots, opportunities and principles of their application]. *Nauchnyj vestnik NGTU – Science bulletin of NSTU*. 2015. vol. 3(60). pp. 147–165. (In Russ.).
4. Prikaz Minzdravsocrazvitija RF № 477n «Ob utverzhdenii perechnja sostojanij, pri kotoryh okazyvaetsja pervaja pomoshh', i perechnja meroprijatij po okazaniju pervoj pomoshhi» [Order of the Health Ministry of the Russian Federation № 477n "On approving the list of conditions for which to provide first aid, and the list of measures for first aid"]. 2012. 3 p. (In Russ.).
5. Miheev M.Ju., Kotjakova V.A., Volodina E.A., Bannov V.Ja. [Application of "decision tree" for the analysis of complex systems]. *Trudy mezhdunarodnogo simpoziuma "Nadjozhnost' i kachestvo"* [Proceedings of international symposium "Reliability & Quality"]. Penza: PGU. 2012. vol. 2. pp. 401–403. (In Russ.).
6. Tutygin A.G., Korobov V.B. [Advantages and disadvantages of the anakytic hierarchy process]. *Izvestija RGPU im. A. I. Gercena – Izvestia: Herzen University Journal of Humanities & Scienc*e. 2010. vol. 1. pp. 108–115. (In Russ.).
7. Produkcionnaja model' znanij [Production system]. Available at: http://www.aiportal.ru/articles/knowledge-models/production-model.html (accessed: 29.10.2015). (In Russ.).
8. Kuz'kin A.A. *Metodika obespechenija ustojchivosti strategii razvitija informacionnyh tehnologij na predprijatii v uslovijah neopredelennosti vozdejstvija sredy* [The methods for ensuring the sustainability of strategy of development of information technology in the enterprise in the conditions of uncertainty of exposure]. Ph.D. thesis. Spb. 2015. 116 p. (In Russ.).
9. Basov O.O., Karpov A.A., Saitov I.A. *Metodologicheskie Osnovy Sinteza Polimodal'nykh Infokommunikatsionnykh Sistem Gosudarstvennogo Upravleniya* [Methodological Bases of Synthesis of Multimodal Communication Systems of Public Administration]. Orel. Akademiya FSO Rossii Publ. 2015. 271 p. (In Russ.).
10. Musina V.F. [Bayesian belief networks as probabilistic graphical model for medical risk assessment]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. vol. 24. pp. 135–151. (In Russ.).

11. Pitsikalis V., Katsamanis A., Papandreou G., Maragos P. Adaptive multimodal fusion by uncertainty compensation. In Proceedings of the Ninth International Conference on Spoken Language Processing. Pittsburgh. 2006.
12. Wasyluk H., Onisko A., Druzdzel M.J. Support of diagnosis of liver disorders based on a causal Bayesian network model. *Medical Science Monitor*. 2001. vol. 7. pp. 327–332.
13. Lacave C., Diez F.J. Knowledge Acquisition in PROSTANET – A Bayesian network for diagnosis prostate cancer. *Knowledge-Based Intelligent Information and Engineering Systems*. 2003. LNCS 2774. pp. 1345–1350.
14. Wiegerincka W.A.J.J., et al. Approximate inference for medical diagnosis. *Pattern Recognition Letters*. 1999. vol. 20. no. 11–13. pp. 1231–1239.
15. Jensen F.V., Nielsen T.D. Bayesian networks and decision graphs. New York: Springer. 2007.
16. Dempster A., Laird N., Rubin D. Maximum likelihood from incomplete data via the EM algorithm. *Jour. of the Royal Statistical Society*. 1997. vol. 39. no. 1. pp. 1–38.
17. Bender J., Koller D., Russel R., Kanazava K. Adaptive probabilistic networks with hidden variables. *Machine Learning*. 1997. vol. 29. no. 2–3. pp. 213–244.
18. Tulup'ev A.L., Sirotkin A.V., Nikolenko S.I. *Bajesovskie seti doverija. Logiko-verojatnostnyj vyvod v aciklicheskih napravlennyh grafah* [Bayesian belief networks. Logical and probabilistic inference in the acyclic directed graph]. SPb.: Izd-vo SPbGU. 2009. 400 p. (In Russ.).
19. Henrion M. Propagating uncertainty in Bayesian networks by logic sampling. *Uncertainty in Artificial Intelligence*. Amsterdam 1988. vol 2. pp. 149–163.
20. Fung R., Chang K.-C. Weighting and integrating evidence for stochastic simulation in Bayesian networks. In Proceedings of the Fifth Conference on Uncertainty in Artificial Intelligence (UAI-89). Amsterdam. 1989. pp. 475–482.
21. Shachter R., Peot M. Simulation approaches to general probabilistic inference on belief networks. In Proceedings of the Fifth Workshop on Uncertainty in Artificial Intelligence (UAI-89). Amsterdam. 1989. pp. 311–318.
22. Maslennikov E.D., Sulimov V.B. [Predictions based on Bayesian belief networks: algorithm and software implementation] *Vychislitel'nye metody i programmirovanie – Numerica methods and programming*. 2010. vol. 11. no. 2. pp. 222–235. (In Russ.).
23. Netica. Available at: http://www.norsys.com/ (accessed: 12.09.2015).
24. Ronzhin A.L., Yusupov R.M. [Multimodal interfaces for autonomous robotic ststems]. *Izvestija JuFU. Tehnicheskie nauki – Izvestiya SFedU. Engineering sciences*. 2015. vol. 1(162). C. 195–206. (In Russ.).
25. *Basov O.O.* [Principles of construction of polymodal info-communication systems based on multimodal architectures of subscriber's terminals]. *Trudy SPIIRAS – SPIIRAS Proceedings*. 2015. vol. 39. pp. 109-122. (In Russ.).
26. Kozyrenko N.K., Meshcheryakov R.V., Hodashinsky I.H., Anufrieva N.A. [Mathematical Model and Algorithms of People Health Evaluation]. *Trudy SPIIRAS – SPIIRAS Proceedings*. 2015. vol. 33. pp. 117-146. (In Russ.).
27. Karpov A.A., Ronzhin A.L. [Multimodal interfaces in automated control systems]. *Izvestija vysshih uchebnyh zavedenij. Priborostroenie – Universities proceedings. Instrument engineering*. 2005. vol. 48. no. 7. pp. 9-14.

Мотиенко Анна Игоревна — научный сотрудник исследовательской группы информационных технологий в образовании, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН), преподаватель кафедры физики, математики и информатики, Первый Санкт-Петербургский государственный медицинский университет им. акад. И. П. Павлова (ГБОУ ВПО ПСПбГМУ им. И. П. Павлова Минздрава России). Область научных интересов: информационный технологии в

образовании, информационные технологии ы медицине, аварийно-спасательные роботы. Число научных публикаций — 4. anna.gunchenko@gmail.com; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7(812)328-03-82.

Motienko Anna Igorevna — researcher of research group of information technologies in education, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), teacher of physics, mathematics and informatics department, Pavlov First Saint Petersburg State Medical University. Research interests: information technologies in education, information technologies in medicine, rescue robots. The number of publications — 4. anna.gunchenko@gmail.com; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)328-03-82.

Макеев Сергей Михайлович — сотрудник, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: оценка техногенного риска, вероятностное и статистическое моделирование, количественные и качественные методы анализа данных. Число научных публикаций — 9. MakSM57@yandex.ru; Приборостроительная, 35, Орел, 302034; р.т.: 8(486)2549464.

Makeev Sergey Mihajlovich — researcher, The Academy of Federal Security Guard Service of the Russian Federation. Research interests: assessment of technogenic risk, probabilistic and statistical modeling, quantitative and qualitative methods of data analysis. The number of publications — 9. MakSM57@yandex.ru; 35, Priborostroitelnaya Street, Orel, 302034, Russia; office phone: 8(486)2549464.

Басов Олег Олегович — к-т техн. наук, сотрудник, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: обработка и кодирование речевых и иконических сигналов, проектирование полимодальных инфокоммуникационных систем. Число научных публикаций — 165. oobasov@mail.ru; Приборостроительная, 35, Орел, 302034; р.т.: +7(4862)549533.

Basov Oleg Olegovich — Ph.D., researcher, The Academy of Federal Security Guard Service of the Russian Federation. Research interests: processing and coding of speech and iconic signals, polymodal infocommunicational systems design. The number of publications — 165. oobasov@mail.ru; 35, Priborostroitelnaya Street, Orel, 302034, Russia; office phone: +7(4862)549533.

РЕФЕРАТ

Мотиенко А.И., Макеев С.М., Басов О.О. **Анализ и моделирование процесса выбора положения для транспортировки пострадавшего на основе байесовских сетей доверия.**

Целью настоящего исследования является разработка методов реализации мероприятий по оказанию первой помощи пострадавшим и программно-алгоритмических решений, обеспечивающих их транспортировку, на основе моделирования автоматизированных робототехнических средств при проведении аварийно-спасательных и других неотложных работ.

Для транспортировки пострадавших с помощью автоматизированных робототехнических средств необходимо определение оптимального положения тела. Ведущую роль при выборе способа, средств и положения, в котором будут транспортироваться пострадавшие, играют виды травм, их локализация, состояние пострадавшего. Однако в рассматриваемой предметной области наблюдается отсутствие адекватного описания соответствий между признаками полученных травм и оптимальным положением тела пострадавшего при транспортировке.

В статье представлена вероятностная модель на основе байесовской сети доверия, которая позволяет определить наиболее важные признаки травм, влияющие на принятие решения о выборе положения пострадавшего при его транспортировке. Вероятностные оценки наблюдаемых свойств (признаков травм) дают фактические величины, позволяющие получить возможность пересмотра доверия к оценкам ненаблюдаемых свойств (травм и положений для транспортировки). В данном исследовании было проведено моделирование процесса выбора положения для транспортировки пострадавшего с использованием программы «Netica». Рассматриваемый подход позволяет определить наиболее важные признаки, влияющие на принятие решения о выборе положения пострадавшего при транспортировке в условиях неполноты данных касательно характеристик травм, состояния человека, а также сформировать список действий, наиболее эффективно приводящих к необходимому заключению о выборе положения для транспортировки.

SUMMARY

Motienko A.I., Makeev S.M., Basov O.O. **Analysis and Modeling of the Process of a Choice of Position for Transportation of the Sufferer on the basis of Bayesian Belief Networks.**

The purpose of this research is the development of methods for the implementation of measures to provide first aid to the sufferer and program-algorithmic solutions, providing the transportation, on the basis of modeling of automated robotic means during the rescue and other emergency operations.

The transportation of the sufferer by automated robotic means needs to determine the optimal position of the body. The leading role in choice of the method, means, and the position in which the sufferer will be transported play types of injuries, their location, condition of the sufferer. However, in the subject area there is a lack of adequate description of correspondences between the signs of the injuries and the optimal position of the body of the sufferer during transportation.

The probabilistic model based on Bayesian belief networks, which allows you to identify the most important signs of injuries affecting the decision on the choice of the position of the sufferer during transportation is presented in the paper. Probabilistic assessments of the observed properties (signs of injures) give the actual values, allowing to get the opportunity to reconsider the credibility of the estimates of unobserved properties (injuries and positions for the transportation). The modeling of process of a choice of position for transportation of the sufferer with use of the Netica program was carried out in this research. The considered approach allows to define the most important signs of injuries influencing the decision on a choice of position of the sufferer during transportation in the conditions of incomplete information about the characteristics of injuries, a condition of the person and also to create a list of actions which most effectively will lead to the necessary conclusion about the choice of position for transportation.

УДК 004.891

А.В. Торопова

ПОДХОДЫ К ДИАГНОСТИКЕ СОГЛАСОВАННОСТИ ДАННЫХ В БАЙЕСОВСКИХ СЕТЯХ ДОВЕРИЯ

Торопова А.В. **Подходы к диагностике согласованности данных в байесовских сетях доверия.**

Аннотация. Байесовские сети доверия предоставляют возможность объединения нескольких видов информации, например полученной от экспертов или статистически, позволяют работать с неполной или неточной информацией, обладают наглядностью и другими полезными свойствами. Благодаря этому они стали популярным и весьма эффективным инструментом. Однако во многих областях исследования исходные используются полученные от экспертов данные, которые могут быть не согласованы, и поэтому в некоторых задачах следует использовать инструменты для проверки их согласованности.

В работе рассмотрены примеры применения аппарата байесовских сетей доверия в медицине и здравоохранении, экологии, экономике и риск-анализе, функциональной безопасности, социологии и других предметных областях и показана необходимость разработки методов для проверки согласованности исходных данных.

Цель работы – систематизировать с помощью обзора примеры и задачи, в которых применяются байесовские сети доверия, чтобы оценить, в какой степени в этих задачах учитывается диагностика согласованности исходных данных, и насколько важным является ее применение.

Ключевые слова: диагностика согласованности данных, байесовские сети доверия.

Toropova A.V. **Approaches to the Data Coherence Diagnosis in Bayesian Belief Network Models.**

Abstract. Bayesian belief networks provide the ability to combine different types of information, e.g. statistical or expert data, allow working with incomplete or inaccurate information; they have clarity and other useful properties. Due to this, Bayesian belief networks have become a popular and highly effective tool in many fields of research. However, in many research areas data provided by the experts can be incoherent, and so in some tasks one should use tools to verify their coherence. The paper discusses examples of application of the Bayesian belief networks in medicine and public health, ecology, economics and risk analysis, functional safety, sociology, and other research areas, and shows the need to develop methods to check the coherence of initial data. The purpose of this work is to systematize problems and examples that illustrate the use of Bayesian belief networks by reviewing and to assess their use of data coherence diagnosis and its importance.

Keywords: data coherence diagnosis, Bayesian belief networks.

1. Введение. Байесовские сети доверия предоставляют возможность объединения нескольких видов информации, например полученной от экспертов или статистически, позволяют работать с неполной или неточной информацией, обладают наглядностью и другими полезными свойствами. Благодаря этому они стали популярным и весьма эффективным инструментом во многих областях исследования.

В работе [11] байесовские сети доверия используются для построения модели оценки интенсивности социально-значимого поведения (это может быть рискованное поведение такое, как потребление

алкоголя или незащищенные половые связи, активность в социальных сетях и др.). Однако в описанной модели возникает проблема диагностики согласованности данных, полученных от респондентов, между собой. Рассмотрим, как подобные вопросы решаются в других предметных областях.

Цель работы – систематизировать с помощью обзора примеры и задачи, в которых применяются байесовские сети доверия, чтобы оценить, в какой степени в этих задачах учитывается диагностика согласованности исходных данных, и насколько важным является ее применение.

Статья построена следующим образом: в разделе 2 приведены некоторые сведения о байесовских сетях доверия, в разделах 3 – 8 рассмотрена проблема диагностики в медицине, экологии, экономике и других областях, и в заключении сделаны выводы.

2. Байесовские сети доверия. Байесовская сеть доверия – это ациклический направленный граф (то есть граф без направленных циклов, допускаются ненаправленные циклы), вершинам которого соответствуют случайные элементы, а ребра между вершинами соответствуют условным зависимостям между элементами [12]. Каждый случайный элемент описывается функцией распределения вероятности, в данном случае – представленной в виде тензора условных вероятностей.

Приведем формальное определение [12].

Пусть X – случайный элемент, принимающий значения из множества $\{x_1,...,x_n\}$. Вероятность $p(X = x_i)$ будем обозначать как $p(x_i)$. Распределение вероятности означивания X будем обозначать как $P(X)$.

Направленный граф $G(V,L)$ – это пара V,L, где V – множество вершин $\{v_1,...,v_n\}$, а L – множество направленных ребер $\{(u,v) \mid u,v \in V, u \neq v\}$. Через $\mathrm{pa}(v)$ будем обозначать множество таких вершин u, что существует ребро (u,v), то есть множество вершин, из которых исходит ребро, направленное в v.

Формально, байесовской сетью доверия выступает пара $\langle G, \mathrm{P} \rangle$, где G – ациклический направленный граф, а P – множество тензоров условных вероятностей $P(X \mid \mathrm{pa}(X))$. Такая пара задает совместное вероятностное распределение над всеми случайными элементами, входящими в сеть, в предположении, что X независим от всех остальных элементов сети при заданном означивании его родителей, детей и родителей детей.

Важной особенностью байесовских сетей доверия как вероятностных графических моделей является правило декомпозиции, которое

обусловлено предположением об условной независимости элементов. Правило декомпозиции выглядит следующим образом: $P(X_1,...,X_m) = P(X_1 \mid \text{pa}(X_1)) \cdot ... \cdot P(X_m \mid \text{pa}(X_m))$, где $P(X_1,...,X_m)$ – совместное вероятностное распределение всех случайных элементов модели, $P(X_i \mid \text{pa}(X_i))$ – вероятностное распределение случайного элемента X_i при условии означивания случайных элементов $\text{pa}(X_i)$ – родителей вершины X_i, $i = 1,...,m$.

Байесовская сеть доверия может быть построена как на основе экспертных оценок, так и на основе статистических данных. Экспертная информация может использоваться для установления взаимосвязей между случайными элементами и для получения оценок условных вероятностей [51]. Правило декомпозиции позволяет использовать алгоритмы вероятностного вывода [11].

3. Медицина и здравоохранение. Байесовские сети за последние пару десятков лет стали очень популярны в биомедицине и здравоохранении благодаря возможности работы с неточными знаниями, которые участвуют в диагностике заболеваний, выборе оптимальной альтернативы и прогнозирования результатов лечения, кроме того, они предлагают очень привлекательное формальное представление неточных знаний (результат объединения статистических методов для анализа данных и инструментов искусственного интеллекта) [16]. Алгебраические байесовские сети имеют логико-вероятностную семантику, и также могут быть использованы в медицинской диагностике [13].

Основные задачи биомедицины и здравоохранения – это диагностика заболеваний, прогнозирование состояния пациентов, выбор подходящего курса лечения и обнаружение функциональных взаимодействий на клеточном уровне. Рассмотрим эти задачи немного подробнее.

3.1. Диагностика заболеваний. Постановка диагноза каждому пациенту представляет собой построение гипотезы о заболевании, от которого страдает пациент, основанное на косвенных наблюдениях и диагностических тестах. Последние, впрочем, не дают стопроцентной ясности о состоянии пациента. Чтобы избежать неправильного диагноза результаты тестов должны быть рассмотрены с учетом построенной гипотезы. Байесовские сети предлагают естественную основу для такого типа рассуждений в условиях неопределенности. Большое число систем были разработаны и разрабатываются на данной основе. Наиболее известная система – это Pathfinder [31].

Формально наиболее вероятный диагноз D^* можно определить как значение из множества возможных диагнозов D, доставляющее

максимум вероятности наличия заболевания при условии конкретного набора свидетельств E , которые включают симптомы, результаты тестов и другие признаки [41]: $D^* = \arg\max_D \Pr(D \mid E)$.

Системы здравоохранения очень сложны и зависят от большого количества организационных, экономических и структурных факторов. Соответствующие им инструменты должны учитывать взаимодействия между элементами, которые определяют поведение таких систем, а также быть понятными в изучении и позволять их анализ с целью улучшения производительности. Поскольку многие из факторов, влияющих на производительность систем здравоохранения неточны, байесовские сети могут сыграть важную роль в их исследовании, в качестве формальной модели для представления знаний и обработки неопределенностей [16].

Байесовские сети доверия позволяют моделировать знания с неопределенностью. Аппарат байесовской сети доверия позволяет комбинировать имеющиеся статистические данные о характеристиках здоровья пациентов в дополнение к экспертной информации, которую предоставляют врачи-специалисты [57]. Кроме того, байесовские сети доверия (по сравнению с другими моделями) позволяют моделировать возможность возникновения нескольких заболеваний [4].

В настоящее время известны случаи применения байесовских сетей доверия для диагностики заболеваний рака груди [23], печени [60], рака яичников [19], рака простаты [35], зубной боли [25] и многих других.

В [23] 25 заболеваний молочной железы (11 злокачественных и 14 доброкачественных) представлены узлами байесовской сети, такие узлы содержат априорные вероятности заболевания (в зависимости от распространенности, возраста, гормональной терапии и других факторов) или возможные выводы по маммографии. Структура модели также состоит из направленных дуг, показывающих условную зависимость элементов сети. Для построения байесовской сети были использованы уже описанные в литературе сведения о зависимостях между переменными. Построенная модель смогла различить доброкачественные и злокачественные заболевания с довольно хорошим результатом (незначительная разница с показаниями специалиста-маммолога).

В [46] описывается система для диагностики слабоумия, состоящая из двух частей: DemNet, которая помогает ЛПР диагностировать вероятность слабоумия, и PathNet, идентифицирующей возможные заболевания, вызвавшие слабоумие.

В описанных задачах используется медицинская экспертная информация, ее особенностью является частичная некорректность и не-

достаточность [29], поэтому диагностика согласованности такой информации могла бы улучшить точность используемых моделей. В [45] предложены методы диагностики согласованности между собой влияния различных заболеваний.

3.2. Прогнозирование. Задача медицинского прогнозирования – определить будущий курс и исходы процессов, связанных с заболеванием [15]. Так как будущее по своей сути неопределенно, медицинское прогнозирование представляет собой рассуждение в условиях неопределенности. Еще одной важной чертой является применение знаний об изменениях в процессах заболеваний со временем. Байесовские сети в целом могут иметь понятную темпоральную структуру (рисунок 1).

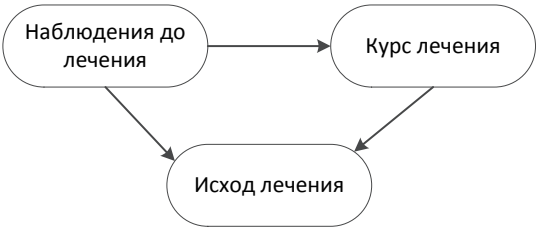


Рис. 1. Общая структура байесовских сетей при прогнозировании [41]

Формально прогноз определяется как вероятность [41]: $\Pr(outcome \mid E, T)$, где E — доступная информация о пациенте (симптомы, результаты обследований, другие признаки), T — выбранный ход лечения. Исходом медицинского вмешательства может служить как ожидаемая продолжительность жизни пациента, так и другие аспекты качества его жизни. В качестве исхода можно рассматривать набор переменных.

Модели медицинского прогнозирования используются как на индивидуальном уровне для поддержки принятия решений о курсе лечения, так и на уровне групп пациентов для управления ресурсами в здравоохранении [58]. Прогностические байесовские сети доверия позволяют обойти трудности, с которыми сталкиваются традиционные прогностические модели. Предсказание в таких моделях обычно представляет собой одношаговый процесс и не предполагает внесения в модель дополнительных переменных, которые становятся известны с течением времени. Кроме того, в общем случае не все переменные, способные повлиять на исход, включаются в модель, например, если два предиктора зависимы, то обычно включается лишь один из них [4]. Известны примеры построения прогностических байесовских сетей в областях онкологии [30] и инфекционных заболеваний [40].

В [16] приведены репрезентационные модели, основанные на байесовских сетях доверия, применимые к конкретному случаю службы экстренной медицинской помощи. Модели построены на основе реальных данных одного испанского госпиталя с использованием алгоритмов обучения байесовских сетей. Эти модели могут быть использованы в задачах управления службами здравоохранения.

В задачах прогнозирования также крайне важно, чтобы исходные данные не имели противоречий, поэтому рекомендуется использовать средства для диагностики их согласованности. Однако в рассмотренных примерах не предложены способы проверки диагностики.

3.3. Выбор лечения. Байесовские сети доверия не предоставляют нужного уровня математической модели для выбора оптимального лечения, однако средства, используемые для решения этой проблемы часто включают байесовские сети доверия в качестве компонент систем поддержки решений для выбора оптимального лечения при данных прогнозирования [18], и в таких компонентах может проводиться диагностика согласованности исходных данных.

Также математическая модель байесовских сетей может быть расширена для включения знаний о решениях и предпочтениях. В качестве примера можно рассмотреть диаграммы влияния [41]. Как и байесовские сети – это ациклический граф. В этом графе множество вершин разделено на множество вершин моделирующих случайные величины, множество вершин, представляющих различные варианты лечения, и вершину, моделирующую предпочтения, которые надо учесть. Структура диаграммы влияния представлена на рисунке 2.

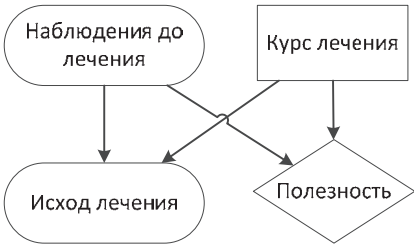


Рис. 2. Общая структура диаграммы влияния [41]

3.4.Обнаружение функциональных взаимодействий. Топологию байесовской сети доверия можно интерпретировать как представление неопределенных взаимодействий между переменными, в связи с этим в биоинформатике возрастает интерес к использованию байесовских сетей доверия для объяснения молекулярных механизмов на клеточном уровне. Например, важная исследовательская область – поиск

взаимодействий между генами на основе экспериментально полученных с помощью микрочипов данных [29].

Биологические данные часто собирают довольно долгое время, анализ временных закономерностей может показать взаимодействие переменных как функцию от времени. Байесовские сети используются для анализа таких биологических временных рядов [48].

В [34] прогнозируется взаимодействие белок – белок в геномных данных дрожжей, но в дальнейшем модель может быть усовершенствована для исследования более сложных организмов.

В [49] показано, как можно использовать байесовские сети в моделировании нейронных сетей.

Также байесовские сети доверия используют в тех областях здравоохранения, которые не относятся напрямую к лечению конкретных пациентов, например, в клинической эпидемиологии их используют для моделирования болезней, а в биоинформатике – для интерпретации данных микрочипов экспрессии генов [41].

Отметим, что в таких задачах диагностика входных данных не приведет к значительным улучшениям, поэтому в данной предметной области не рассматриваются методы проверки согласованности.

4. Экология. Также байесовские сети доверия часто используются в области экологии [22]. Байесовские сети доверия – это эффективное средство для структурирования экологических исследований. Есть два основных пути их применения [42]. Первый заключается в оценке понимания функционирования исследуемых экосистем. В таком случае исследование фокусируется на связях байесовских сетей доверия и обращается к функциональным взаимоотношениям в экосистеме или на «правилах», используемых для построения условных вероятностей для узла, и обращается к механизмам, описывающим взаимодействие факторов в определении значений переменных. Моделирование байесовских сетей доверия в основном касается таких вопросов, как «какие экологические процессы вовлечены?», «какие наиболее важны в воздействии на результаты?», «как они взаимодействуют и как прогнозирование значений переменных может оказаться полезным в экологических процессах?»

Второй путь состоит в использовании байесовских сетей доверия в оценке значений переменных, представленных узлами. Тогда исследование фокусируется на оценках, проверяющих модель и предоставляющих эмпирическую информацию, которая является количественной, полезной и относящейся к ключевым экологическим переменным. Байесовские сети доверия помогают определить сильно влияющие на результаты, но еще не очень хорошо изученные, пере-

менные, поддерживая при этом структурирование и создание средств для тестирования ответной реакции на управляющие решения.

В [28] байесовские сети доверия используются для моделирования и прогнозирования неисправностей в системе распределения питьевой воды.

В [55] – для создания модели оценки рисков угроз почвы. Чтобы определить риск уплотнения почвы, требуются или данные о поведении почвы, собрать которые требует больших затрат, или экспертные данные, которые зачастую имеют субъективный характер. В [55] предлагается модель на основе байесовских сетей доверия, объединяющая данные, полученные стандартными тестами почв, и качественные экспертные данные. В модели три результирующих вершины, дающих представление о том, насколько подвержена почва уплотнению, и двадцать девять вершин, представляющих различные характеристики почвы (текстура, содержание в почве различных веществ и др.), климатические факторы (влажность, температура, время года и др.) и другие факторы, воздействующие на почву (как используется исследуемая территория, как обрабатывается и др.), значения некоторых из этих вершин предоставляется экспертами, однако о диагностике согласованности этих данных информация отсутствует.

В [14] предложено применение куста событий и байесовской сети доверия для оценки экологической ситуации в зоне влияния потенциально химически опасных объектов и вероятности тех или иных ситуаций, связанных с его функционированием. В предложенной модели [14] содержатся вершины, отвечающие за ландшафтные условия, сезонность, время суток, метеоусловия, технологические и технических характеристики работы потенциально химически опасных объектов, условия хранения, объем и параметры загрязняющих веществ, воздействие загрязняющих веществ на персонал, население и окружающую среду, оценка экологической безопасности объекта и связи между ними (рисунок 3). О том, каким образом получены и анализируются ли исходные данные в работе не указано.

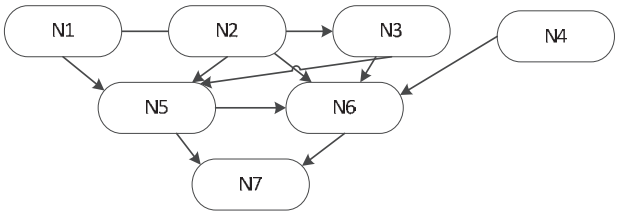


Рис. 3. Модель байесовской сети доверия: N1, N2 – ландшафтные условия, сезонность, время суток, метеоусловия; N3, N4 – характеристики работы ПХОО, условия хранения, объем и параметры загрязняющих веществ;N5, N6 – воздействие загрязняющих веществ; N7– оценка экологической безопасности объекта [14]

Байесовские сети доверия являются хорошим инструментом для оценивания работы экосистемных служб. Так, в работе [36] они использованы для оценки работы экосистемных служб управления озерным комплексом в Бельгии.

Так как в этих задачах в основном используется смесь экспертных данных и точных технических показаний, инструменты диагностирования экспертных данных могут помочь достичь более точных результатов.

5. Функциональная безопасность. Еще одна важная сфера применения аппарата байесовских сетей доверия – функциональная безопасность. Функциональная безопасность – одна из важнейших частей многих современных технических систем. Она включает в себя процессы проектирования и оперативных мер, критически важных для систем безопасности [32]. Байесовские сети доверия используются в диагностике различных систем. В [9] описаны методы для построения экспертных систем для задач технической диагностики. В [3] предлагается методика диагностирования сложных систем, разработанная на основе апостериорного вывода в байесовских сетях доверия. В [62] описан подход к обнаружению элементов систем, вызывающих проблемы с производительностью. В [61] рассматриваются вопросы диагностики и мониторинга систем энергоснабжения. В [43] исследуется диагностика неисправностей в системах электроснабжения, в частности в бортовых системах управления. В [51] предложен метод решения проблемы диагностики газового тракта реактивных двигателей. Модель, предложенную в [38], можно использовать для диагностики производственных процессов (в качестве примера рассматривается литейная промышленность).

В [8] исследуется функционирование энергетических отраслей в условиях чрезвычайных ситуаций. Рассмотрен пример функционирования энергетической системы в ситуации похолодания: на первом этапе исследования разрабатывается онтология, описывающая взаимосвязи основных факторов в рамках угрозы похолодания; на ее основе строится БСД-модель угрозы похолодания, после этого строится событийная карта угрозы похолодания, и заключительным этапом является проведение численных расчетов на сформированных информационных моделях.

В [59] предложен графический подход для диагностики валидности предположений об условной независимости данных байесовской сети. Авторы сфокусировались на определении условной независимости и разработали графическую диагностику, которая показывает поддержанные данными предположения о независимости в структуре байесовской сети, разработали и описали метод Монте–Карло для генерации измерений неточности для согласованности данных с предположениями

об их условной независимости, то есть с помощью этого метода можно графически отобразить являются ли величины, отображающиеся вершинами байесовской сети условно независимыми или нет.

В системах, исследующих функциональную безопасность, исходные данные обычно организованы таким образом, что диагностика их согласованности не требуется.

6. Экономика и риск-анализ. Байесовские сети доверия применяются в задачах оценки экономических рисков и поддержки принятия решений в условиях неопределенности в контексте риск-менеджмента предприятий [5].

Байесовские сети доверия представляют собой удобный аппарат для анализа рисков и их количественной оценки с учетом максимального числа факторов, позволяющий кроме всего прочего наглядно отобразить структуру системы с точки зрения риск-анализа. Байесовские сети позволяют использовать как вероятности полученные аналитическим или статистическим путем, так и экспертные оценки [10].

Значение вероятности некоторого фактора в байесовской сети доверия определено изменениями в других связанных факторах. Механизм вывода основан на теореме Байеса, что делает возможным вычислить вероятность эффекта на какую-либо переменную в модели в зависимости от заданной причины. В случае с двумя связанными переменными вероятности могут быть посчитаны следующим образом:

$$P[эффект] = [P[эффект / причина] \cdot P[причина]] \; / \; P[причина / эффект],$$

где $P[причина]$ – вероятность того, что причина произойдет, $P[эффект]$ – вероятность того, что эффект возникнет, $P[эффект \; / \; причина]$ – условная вероятность эффекта при том, что причина произойдет, $P[причина \; / \; эффект]$ – условная вероятность причины, учитывая, что эффект возникнет.

Апостериорная вероятность причины, при том что эффект возникнет, может быть выведена следующим образом:

$$P[причина \; / \; эффект] = (P[эффект \; / \; причина] * P[причина]) / P[эффект].$$

Байесовские сети доверия используются для построения моделей рисков, состоящих из сценариев, основываясь на множестве известных возможных факторов риска, связанных с анализируемыми рисками. Возможные сценарии структурируются как множество взаимоисключающих и исчерпывающих элементов, которым может быть присвоено вероятностное распределение [24]. В [26] на основе байесовских сетей доверия с использованием обратной связи (feedback loop) была создана процедура для прогнозирования рисков и обнаружения их источников.

В статье [21] описан процесс построения байесовской сети доверия на основании данных, содержащих лишь маргинальные вероятности реализации риска и корреляции между различными типами рисков. В основе предложенного метода лежит решение систем уравнений, которые позволяют вычислить условные вероятности по имеющимся данным.

Вопросы анализа и моделирования операционных рисков, то есть рисков убытка в результате ошибочных внутренних процессов, действий сотрудников и систем или внешних событий, исследуются в [17]. Байесовские сети доверия позволяют корректно конструировать распределение операционных рисков в тех случаях, когда реализации рисков зависимы во времени [20] или в случаях, когда частота реализации риска и последствия реализации зависимы [7].

Информационные риски, то есть возможные угрозы информационной безопасности в любой области деятельности человека [10], также могут анализироваться с помощью байесовских сетей доверия.

В работе [37] аппарат байесовских сетей доверия использован для оценки эффективности работы персонала и организационного здоровья предприятия. В [39] показано, что при использовании этого аппарата можно не только качественно моделировать причинно-следственных связи между организационными факторами и надежности человека, но и количественно измерить эксплуатационную надежность человека, а также определить наиболее вероятные первопричины человеческих ошибок. В [44] предлагается модель для оценки факторов влияющих на бдительность рабочего, следящего и предупреждающего о приближении поездов. В [1] решалась задача оценки вероятности безошибочной работы человека-оператора, управляющего сложным технологическим объектом, тем самым определения его профессиональной надежности, как решения части общей проблемы оценки надежности человеко-машинных систем.

В [56] байесовские сети доверия используются в риск-анализе в области морского судоходства. В [6] описан метод оценки навигационной безопасности при плавании по внутренним водным путям. В [53] – при моделировании рисков природных угроз. В [2] анализируются инновационные риски, предлагается пример структуры байесовской сети для анализа среды на величину риска реализации инновационного проекта.

В задачах экономики и риск-анализа частым является интеграция различной информации, в частности статистических и полученных от экспертов данных. Данные полученные от экспертов могут содержать противоречия, поэтому следует иметь инструмент, помогающий определить насколько согласованы такие данные.

7. Социологические исследования. Применение байесовских сетей доверия в социологических исследованиях сталкивается со следующей проблемой: так как основные сведения для исследований предоставляются различными группами исследуемых, и с большой вероятностью такие данные могут быть не согласованы между собой, необходима проверка их согласованности, то есть оценка того, в какой степени исследователь может им доверять.

В [11] предложена модель для оценки интенсивности социально-значимого поведения, однако в модели не учитывается согласованность исходных данных, что может давать неточные результаты при ее исследовании. Модель представляет собой байесовскую сеть доверия с вершинами, характеризующими интенсивность поведения (*Rate*) длины интервалов между последними тремя эпизодами поведения (*t01*, *t12*, *t23*, *t_min*, *t_max*) и количество дней в исследуемом периоде (*n*) (рисунок 4).

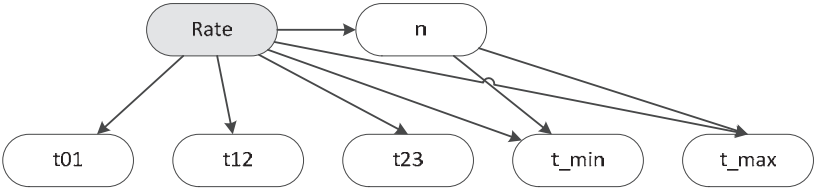


Рис. 4. Модель рискованного поведения, основанная на данных об эпизодах поведения [11]

Единственная проверка данных полученных от респондентов заключается в сравнении максимального и минимального интервалов, однако так как респонденты могут ошибаться или даже заведомо давать ложные сведения, требуется диагностика их согласованности.

Байесовские сети применяются также в анализе социальных сетей [33], в задачах, где учитываются связи между членами сетей, диагностика согласованности данных не нужна.

8. Другие области применения. Байесовские сети доверия используются и в других областях. Например, в астрономии: в [50] их используют как классификатор и средство интегрирования и исследования потоков данных, полученных с помощью специального телескопа (CTI-II). В архитектуре: в [54] представлена модель Architecture Rationale and Element Linkage (AREL) (Рациональная Архитектура и связь элементов), которая представляет причинные взаимосвязи между архитектурными элементами и решениями в модели архитектурного дизайна, что позволяет архитекторам количественно предсказывать и диагностировать влияние изменения части требований или дизайна. В

области разработки программного обеспечения байесовские сети доверия также находят применение. Они могут быть использованы для оценки стоимости программного обеспечения на этапе его разработки [47], для обнаружения его дефектов, оставшихся необнаруженными во время тестирования [27].Задача диагностики согласованности входных данных в этих областях не была затронута.

9. Заключение. Байесовские сети доверия – это удобный и эффективный инструмент в тех случаях, когда возникает необходимость комбинировать разные виды информации, в частности экспертную и статистическую. Экспертные данные, довольно часто использующиеся при сборе информации, а также в моделях принятия решений (вследствие того, что полная математическая формализация многих задач бывает невозможна из-за большой сложности), могут быть не согласованы, а в некоторых случаях и недостоверны. Поэтому актуальна задача разработки инструментария, позволяющего проводить проверку их согласованности. Первым этапом на пути к этому является исследование областей использования байесовских сетей доверия.

В данной работе приведены примеры применения аппарата байесовских сетей доверия в разных предметных областях. Большой популярностью они пользуются в частности в медицине, экологии, экономике, социологии. Указано, что в тех случаях, когда используются экспертные сведения, например в задачах медицинской диагностики и прогнозирования, риск-анализа и др., диагностика согласованности исходных данных особенно необходима. Такая диагностика позволит улучшить модель, сделать ее более надежной, оценить полученный результат. В связи с этим нужно предоставлять возможность оценивания согласованности данных в моделях, что требует разработки или адаптации соответствующих математических и программных методов.

В задачах функциональной безопасности, обнаружения функциональных взаимодействий, анализа социальных сетей данные обычно организованы таким образом, что диагностика согласованности входных данных не требуется.

В дальнейшем планируется продолжение работы над моделью, предложенной в [11], а именно решение проблемы диагностики согласованности данных в ней. Байесовские сети доверия позволяют расширять исследуемые модели новыми узлами и связями и подходят, в том числе и для оценивания согласованности данных, поэтому предлагается дополнение модели диагностическими узлами и исследование эффективности такого метода. По аналогии с этим можно будет создавать инструментарий для анализа согласованности и в более общем случае.

Литература

1. *Бабиков В.М.* Некоторые аспекты применения байесовых сетей для оценки надежности автоматизируемых человеко-машинных систем // Труды международной научно-практической конференции «Передовые информационные технологии, средства и системы автоматизации и их внедрение на российских предприятиях» АІТА-2011. Москва. 2011. С. 266–276.
2. *Белозерский А.Ю., Какатунова Т.В., Иванова И.В.* Использование аппарата нечетких байесовых сетей для оценки инновационных рисков // Транспортное дело России. 2011. № 2. С. 43–46.
3. *Дорожко И.В., Осипов Н.А.* Методика синтеза оптимальных стратегий диагностирования автоматизированных систем управления сложными техническими объектами с использованием априорной информации // Труды СПИИРАН. 2012. № 1(20). С. 165–185.
4. *Мусина В.Ф.* Байесовские сети доверия как вероятностная графическая модель для оценки медицинских рисков // Труды СПИИРАН. 2013. Вып. 24. С. 135–151.
5. *Мусина В.Ф.* Байесовские сети доверия как вероятностная графическая модель для оценки экономических рисков // Труды СПИИРАН. 2013. Вып. 25. С. 235–254.
6. *Некрасов С.Н., Прохоренков А.А.* Комбинированный метод оценки навигационной безопасности при плавании по внутренним водным путям // Вестник государственного университета морского и речного флота им. адмирала С.О. Макарова. 2011. № 1. С. 106–108.
7. *Николаева М.А., Зотова О.Ф., Агадуллина А.И.* Модели и методы управления рисками в социально-экономических системах // Управление риском. 2013. № 1(65). С. 28–34.
8. *Пяткова Е.В.* Технология комплексных исследований функционирования энергетических отраслей в условиях чрезвычайных ситуаций с применением байесовских сетей // Наука и образование: электронное научно-техническое издание. 2013. № 8. С. 293–314.
9. *Солодкий Е.М., Петроченков А.Б.* Экспертная оценка в задачах технической диагностики // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. 2009. № 3. С. 209–215.
10. *Студенников К.О., Лопин В.Н.* О моделировании рисков информационных систем на основе байесовских сетей // Вопросы защиты информации. 2013. № 4(102). С. 3–8.
11. *Суворова А.В.* Модели и алгоритмы анализа сверхкоротких гранулярных временных рядов на основе байесовских сетей доверия // Дисс. к. ф.-м. н. 2013.
12. *Тулупьев А. Л., Сироткин А. В., Николенко С. И.* Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах // СПб.: Изд-во С.-Петерб. ун-та. 2009. 400 с.
13. *Фильченков А.А.* Алгебраическая байесовская сеть как основа для медицинской диагностической модели // «Математическое и компьютерное моделирование в биологии и химии. Перспективы развития». Сборник трудов I Международной интернет-конференции. Казань: Из-во «Казанский университет». 2012. С. 162–166.
14. *Янников И.М., Телегина М.В., Габричидзе Т.Г.* Оценка экологической ситуации с применением методов математического моделирования // Вектор науки Тольяттинского государственного университета. 2011. № 4. С. 38–41.
15. *Abu-Hanna A., Lucas P. J F.* Prognostic Models in Medicine: AI and Statistical Approaches // Methods of Information in Medicine. 2001. vol. 40. pp. 1–5.
16. *Acid S., de Campos L.M., Fernández-Luna J.M., Rodríguez S., et al.* A comparison of learning algorithms for Bayesian networks: a case study based on data from an emer-

gency medical service // Artificial Intelligence in Medicine. 2004. vol. 30. no. 3. pp. 215–232.
17. *Alexander C.* Managing operational risks with Bayesian networks // Operational Risk: Regulation, Analysis and Management. 2003. pp. 285–294.
18. *Andreassen S., et al.* Using probabilistic and decisiontheoretic methods in treatment and prognosis modeling // Artif Intell Med. 1999. vol. 15. pp. 121–134.
19. *Antal P., Verrelst H., Timmerman D., Moreau Y., et al.* Bayesian Networks in Ovarian Cancer Diagnosis: Potentials and Limitations // Proceedings of the 13th IEEE Symposium on Computer-Based Medical Systems (CBMS 2000). 2000. pp. 103–108.
20. *Aquaro V., Bardoscia M., Bellotti R., Consiglio A., et al.* A Bayesian Networks approach to Operational Risk // Physica A: Statistical Mechanics and its Applications. 2010. vol. 389. no. 8. pp. 1721–1728.
21. *Bonafede C.E., Giudici P.* Bayesian networks for enterprise risk assessment // Physica A: Statistical Mechanics and its Applications. 2007. vol. 382. Issue 1. pp. 22–28.
22. *Borsuk M.E., Stow C.A., Reckhow K.H.* A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis // Ecological Modelling. 2004. vol. 173. no 2. pp. 219–239.
23. *Burnside E.S., Rubin D.L., Fine J.P., Shachter R.D., et al.* Bayesian Network to Predict Breast Cancer Risk of Mammographic Microcalcifications and Reduce Number of Benign Biopsy Results: Initial Experience // Radiology. 2006. vol. 240. no. 3. pp. 666–673.
24. *Cárdenas I.C., Al-jibouri S.SH., Halman J. IM.* A Bayesian Belief Networks Approach to Risk Control in Construction Projects // 14th International Conference on Computing and Civil Engineering. Moscow. Russian Federation. The ISCCBE and Moscow State University of Civil Engineering (National Research University).
25. *Chattopadhyay S., Davis R.M., Menezes D.D., Singh G., et al.* Application of Bayesian Classifier for the Diagnosis of Dental Pain // J Med Syst. 2012. vol. 36. pp. 1425–1439.
26. *Fan C.F., Yu Y.C.* BBN-based software project risk management // Journal of Systems and Software. 2004. vol. 73. no 2. pp. 193–203.
27. *Fenton N., Neil M., Marsh W., Hearty P., et al.* Predicting software defects in varying development lifecycles using Bayesian nets // Information and Software Technology. 2007. vol. 49. no. 1. pp. 32–43.
28. *Francis R.A., Guikema S.D., Henneman L.* Bayesian belief networks for predicting drinking water distribution system pipe breaks // Reliability Engineering & System Safety. 2014. vol. 130. pp. 1–11.
29. *Friedman N.I.R., Linial M., Nachman I., Pe'er D.* Using Bayesian network to analyze expression data // J Comput Biol. 2000. vol. 7. pp. 601–20.
30. *Galan S.F., Aguado F., Diez F.J., Mira J.* NasoNet: joining Bayesian networks and time to model nasopharyngeal cancer spread // Proceedings of the Eighth International Conference on Artificial Intelligence in Medicine in Europe (AIME 2001). Berlin: Springer-Verlag. 2001. LNAI 2101. pp. 207–216.
31. *Heckerman D.E., Horvitz E.J., Nathwani B.N.* Towards normative expert systems. Part I. The Pathfinder project // Methods of information in medicine. 1992. vol. 31. pp. 90–105.
32. *Hofbaur M., Sachenbacher M.* On the Role of Model-based Diagnosis in Functional Safety // Proceedings of the 24th International Workshop on Principles of Diagnosis. Jerusalem. 2013. pp. 65–70.
33. *Koelle D., Pfautz J., Farry M., Cox Z., et al.* Applications of bayesian belief networks in social network analysis // Proceedings of the 4th Bayesian Modeling Applications Workshop during the 22nd Annual Conference on Uncertainty in Artificial Intelligence. 2006.

34. *Jansen R., Yu H., Greenbaum D., Kluger Y., et al.* A bayesian networks approach for predicting protein-protein interactions from genomic data // Science. 2003. vol. 302(5644). pp. 449–453.
35. *Lacave C., Diez F.J.* Knowledge Acquisition in PROSTANET – A Bayesian network for diagnosis prostate cancer // Knowledge-Based Intelligent Information and Engineering Systems. Lecture Notes in Computer Science. 2003. vol. 2774. pp. 1345–1350.
36. *Landuyt D., Lemmens P., D'hondt R., Broekx S., et al.* An ecosystem service approach to support integrated pond management: A case study using Bayesian belief networks — Highlighting opportunities and risks // Journal of Environmental Management. 2014. vol. 145. pp. 79–87.
37. *Léger A., Duval C., Farret R., Weber P., et al.* Modeling of human and organizational impacts for system risk analyses // 9th International Probabilistic Safety Assessment and Management Conference. Hong Kong. 2008.
38. *Lewis R.W., Ransing R.S.* A semantically constrained Bayesian network for manufacturing diagnosis // International Journal of Production Research. 1997. vol. 35. no. 8. pp. 2171–2188.
39. *Li P.C., Chen G.H., Dai L.C., Zhang L.* A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRA frameworks // Safety Science. 2012. vol. 50(7). pp. 1569–1583.
40. *Lucas P.J.F., De Bruijn N.C., Schurink K., Hoepelman I.M.* A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU // ArtifIntell Med. 2000. vol. 19(3). pp. 251–279.
41. *Lucas P., van der Gaag L., Abu-Hanna A.* Bayesian networks in biomedicine and healthcare // Artificial Intelligence in Medicine. 2004. vol. 30. pp. 201–214.
42. *McCann R. K., Marcot B. G., Ellis R.* Bayesian belief networks: applications in ecology and natural resource management // Canadian Journal of Forest Research. 2006. vol. 36. no. 12. pp. 3053–3062.
43. *Mengshoel O.J., Darwiche A., Cascio K., Chavira M., et al.* Diagnosing faults in electrical power systems of spacecraft and aircraft // Proceedings of IAAI-08. Chicago. 2008. pp. 1699–1705.
44. *Molloy B., Balfe N., Lowe E.* Developing Bayesian Belief Networks to Support Risk-Based Decision Making in Railway Operations // Conference Proceedings: Applied Human Factors and Ergonomics. Poland. 2014.
45. *Nikovsky D.* Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics // IEEE Transactions on Knowledge and Data Engineering. 2000. vol. 12. no. 4. pp. 509–516.
46. *Oteniya L., Cowie J., Coles R.* Diagnosis of Dementia and its Pathologies Using Bayesian Belief Networks // Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration (ICEIS 2006). Cyprus. 2006.
47. *Radlinski L.* A survey of bayesian net models for software development effort prediction // International Journal of Software Engineering and Computing. 2010. vol. 2. no. 2. pp. 95–109.
48. *Ramoni M., Sebastiani P., Cohen P.* Bayesian clustering by dynamics // Mach Learn 2002. vol. 47. pp. 91–121.
49. *Rao Rajesh P.N.* Neural Models of Bayesian Belief Propagation // The Bayesian Brain: Probabilistic Approaches to Neural Coding. Cambridge. MIT Press. 2007. 239 p.
50. *Ritthaler M., Luger G., Young R.* Bayesian Belief Networks for Astronomical Object Recognition and Classification in CTI-II // Astronomical Data Analysis Software and Systems XVI.ASP Conference Series. 2007. vol. 376. pp. 413–416.
51. *Romessis C., Mathioudakis K.* Bayesian network approach for gas path fault diagnosis // Journal of engineering for gas turbines and power. 2006. vol. 128(1). pp. 64–72.

52. *Spiegelhalter D.J., Dawid A.P., Lauritzen S.L., Cowell R.G.* Bayesian Analysis in Expert Systems // Statistical Science. 1993. vol. 8. no. 3. pp. 219–247. URL: http://www.jstor.org/stable/2245959 (дата доступа: 02.11.2015).
53. *Straub D.* Natural hazards risk assessment using Bayesian networks // Safety and Reliability of Engineering Systems and Structures. 2005. pp. 2535–2542.
54. *Tang A., Nicholson A., Jin Y., Han J.* Using Bayesian belief networks for change impact analysis in architecture design // Journal of Systems and Software. 2007. vol. 80. no. 1. pp. 127–148.
55. *Troldborg M., Aalders I., Towers W., Hallett P.D., et al.* Application of Bayesian Belief Networks to quantify and map areas at risk to soil threats: Using soil compaction as an example // Soil and Tillage Research. 2013. vol. 132. pp. 56–68.
56. *Trucco P., Cagno E, Ruggeri F., Grande O.* A Bayesian Belief Network modeling of organizational factors in risk analysis: a case study in maritime transportation // Reliability Engineering and System Safety. 2008. vol. 93. pp. 823–834.
57. *Twardy C.R., Nicholson A.E., Korb K.B., McNeil J.* Epidemiological data mining of cardiovascular Bayesian networks // Electronic Journal of Health Informatics. 2006. vol. 1(1). pp. 3.
58. *Verduijn M., et al.* Prognostic Bayesian networks I: rationale, learning procedure, and clinical use // Journal of Biomedical Informatics. 2007. vol. 40(6). pp. 609–618.
59. *Walsh S., Whitney P.* A Graphical Approach to Diagnosing the Validity of the Conditional Independence Assumptions of a Bayesian Network Given Data // Journal of Computational and Graphical Statistics. 2012. vol. 21. no. 4. pp. 961–978.
60. *Wasyluk H., Onisko A., Druzdzel M.J.* Support of diagnosis of liver disorders based on a causal Bayesian network model // Medical Science Monitor. 2001. vol. 7. no. 1. pp. 327–332.
61. *Yongli Z., Limin H., Jinling L.* Bayesian network-based approach for power system fault diagnosis // IEEE Transactions on Power Delivery. 2006. vol. 21. pp. 634–639.
62. *Zhang R., Moyle S., McKeever S., Bivens A.* Performance problem localization in selfhealing, service-oriented systems using bayesian networks // In Proceedings of the 2007 ACM symposium on Applied computing (SAC '07). 2007.

References

1. Babikov V.M. [Some aspects of the application of Bayesian networks to assess the reliability of the automated man-machine systems]. *Trudy mezhdunarodnoi nauchno-prakticheskoi konferentcii «Peredovye informatcionnye tekhnologii, sredstva i sistemy avtomatizatcii i ikh vnedrenie na rossiiskikh predpriiatiiakh» AITA-2011* [Proceedings of the International scientific and practical conference "Advanced information technologies, tools and automation systems and their implementation at Russian enterprises» AITA-2011]. Moscow. 2011. pp. 266–276. (In Russ.).
2. Belozerskii A.Iu., Kakatunova T.V., Ivanova I.V. [The use of fuzzy bayesian networks to assess the risks of innovation]. *Transportnoe delo Rossii – Transport business in Russia*. 2011. vol. 2. pp. 43–46. (In Russ.).
3. Dorozhko I.V., Osipov N.A. [Technique of synthesis of optimal strategies for diagnostics of the automated control systems of complex technical objects with the use of aprioristic information]. *Trudy SPIIRAN – SPIIRAS Proceedings.* 2012. vol. 1(20). pp. 165–185. (In Russ.).
4. Musina V.F. [Bayesian belief networks as probabilistic graphical model for medical risk assessment]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. vol. 24. pp. 135–151. (In Russ.).
5. Musina V.F. [Bayesian belief networks as probabilistic graphical model for economical risk assessment]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. vol. 25. pp. 235–254. (In Russ.).

6. Nekrasov S.N., Prohorenkov A.A. [Combined simulation while estimation of inland waterway navigation safety]. *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota im. admirala S.O. Makarova – Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S.O. Makarova*. 2011. vol. 1. pp. 106–108. (In Russ.).
7. Nicolaeva M.A., Zotova O.F., Agadullina A.I. [Models and methods for risk management in the social and economic systems]. *Upravlenie riskom – Risk Management*. 2013. vol. 1(65). pp. 28–34. (In Russ.).
8. Piatkova E.V. [Technology of complex research of energy sector functioning in emergency situations with the use of Bayseian networks]. *Nauka i obrazovanie: elektronnoe nauchno-tekhnicheskoe izdanie – Science and Education: electronic scientific and technical journal*. 2013. vol. 8. pp. 293–314. (In Russ.).
9. Solodkii E.M., Petrochenkov A.B. [Peer review in the problems of technical diagnostics]. *Vestnik Permskogo natcionalnogo issledovatelskogo politekhnicheskogo universiteta. Elektrotekhnika, informatcionnye tekhnologii, sistemy upravleniia – Vestnik of Perm National Research Polytechnic University. Electrical engineering, information technology, control systems*. 2009. vol. 3. pp. 209–215. (In Russ.).
10. Studennikov K.O., Lopin V.N. [About information systems risk modeling based on bayesian networks]. *Voprosy zashchity informatcii – Information Security issues*. 2013. vol. 4(102). pp. 3–8. (In Russ.).
11. Suvorova A.V. *Modeli i algoritmy analiza sverkhkorotkikh granuliarnykh vremennykh riadov na osnove baiesovskikh setei doveriia* [Models and algorithms for the analysis of granular ultrashort time series based on Bayesian belief networks]. Ph.D. thesis. 2013. (In Russ.).
12. Tulupev A. L., Sirotkin A. V., Nikolenko S. I. *Baiesovskie seti doveriia: logiko-veroiatnostnyi vyvod v atciclicheskikh napravlennykh grafakh* [Bayesian belief networks: the logic-probabilistic inference in the acyclic directed graph]. SPb.: Izd-vo S.-Peterb. un-ta, 2009. 400 p. (In Russ.).
13. Filchenkov A.A. [Algebraic Bayesian network as the basis for medical diagnostic model]. *«Matematicheskoe i kompiuternoe modelirovanie v biologii i himii. Perspektivy razvitiia». Sbornik trudov I Mezhdunarodnoi internet-konferentcii* ["Mathematical and computer modeling in biology and chemistry. Development prospects". Proceedings of the I International Internet Conference]. Kazan. «KSU ». 2012. pp. 162–166. (In Russ.).
14. Iannikov I.M., Telegina M.V., Gabrichidze T.G. [Estimation of the ecological situation with application of methods of mathematical modelling]. *Vektor nauki Toliattinskogo gosudarstvennogo universiteta – Vector of sciences*. Togliatti State University. 2011. no. 4. pp. 38–41. (In Russ.).
15. Abu-Hanna A., Lucas P. J F. Prognostic Models in Medicine: AI and Statistical Approaches. *Methods of Information in Medicine*. 2001. vol. 40. pp. 1–5.
16. Acid S., de Campos L.M., Fernández-Luna J.M., Rodríguez S., et al. A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service. *Artificial Intelligence in Medicine*. 2004. vol. 30. no. 3. pp. 215–232.
17. Alexander C. Managing operational risks with Bayesian networks. *Operational Risk: Regulation, Analysis and Management*. 2003. pp. 285–294.
18. Andreassen S., et al. Using probabilistic and decisiontheoretic methods in treatment and prognosis modeling. *Artif Intell Med*. 1999. vol. 15. pp. 121–134.
19. Antal P., Verrelst H., Timmerman D., Moreau Y., et al. Bayesian Networks in Ovarian Cancer Diagnosis: Potentials and Limitations. Proceedings of the 13th IEEE Symposium on Computer-Based Medical Systems (CBMS 2000). 2000. pp. 103–108.

20. Aquaro V., Bardoscia M., Bellotti R., Consiglio A., et al. A Bayesian Networks approach to Operational Risk. *Physica A: Statistical Mechanics and its Applications*. 2010. vol. 389. no. 8. pp. 1721–1728.
21. Bonafede C.E., Giudici P. Bayesian networks for enterprise risk assessment. *Physica A: Statistical Mechanics and its Applications*. 2007. vol. 382. Issue 1. pp. 22–28.
22. Borsuk M.E., Stow C.A., Reckhow K.H. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecological Modelling*. 2004. vol. 173. no 2. pp. 219–239.
23. Burnside E.S., Rubin D.L., Fine J.P., Shachter R.D., et al. Bayesian Network to Predict Breast Cancer Risk of Mammographic Microcalcifications and Reduce Number of Benign Biopsy Results: Initial Experience. *Radiology*. 2006. vol. 240. no. 3. pp. 666–673.
24. Cárdenas I.C., Al-jibouri S.SH., Halman J. IM. A Bayesian Belief Networks Approach to Risk Control in Construction Projects. 14th International Conference on Computing and Civil Engineering. Moscow. Russian Federation. The ISCCBE and Moscow State University of Civil Engineering (National Research University).
25. Chattopadhyay S., Davis R.M., Menezes D.D., Singh G., et al. Application of Bayesian Classifier for the Diagnosis of Dental Pain. *J Med Syst*. 2012. vol. 36. pp. 1425–1439.
26. Fan C.F., Yu Y.C. BBN-based software project risk management. *Journal of Systems and Software*. 2004. vol. 73. no 2. pp. 193–203.
27. Fenton N., Neil M., Marsh W., Hearty P., et al. Predicting software defects in varying development lifecycles using Bayesian nets. *Information and Software Technology*. 2007. vol. 49. no. 1. pp. 32–43.
28. Francis R.A., Guikema S.D., Henneman L. Bayesian belief networks for predicting drinking water distribution system pipe breaks. *Reliability Engineering & System Safety*. 2014. vol. 130. pp. 1–11.
29. Friedman N.I.R., Linial M., Nachman I., Pe'er D. Using Bayesian network to analyze expression data. *J Comput Biol*. 2000. vol. 7. pp. 601–20.
30. Galan S.F., Aguado F., Dıez F.J., Mira J. NasoNet: joining Bayesian networks and time to model nasopharyngeal cancer spread. Proceedings of the Eighth International Conference on Artificial Intelligence in Medicine in Europe (AIME 2001). Berlin: Springer-Verlag. 2001. LNAI 2101. pp. 207–216.
31. Heckerman D.E., Horvitz E.J., Nathwani B.N. Towards normative expert systems. Part I. The Pathfinder project. *Methods of information in medicine*. 1992. vol. 31. pp. 90–105.
32. Hofbaur M., Sachenbacher M. On the Role of Model-based Diagnosis in Functional Safety. Proceedings of the 24th International Workshop on Principles of Diagnosis. Jerusalem. 2013. pp. 65–70.
33. Koelle D., Pfautz J., Farry M., Cox Z., et al. Applications of bayesian belief networks in social network analysis. Proceedings of the 4th Bayesian Modeling Applications Workshop during the 22nd Annual Conference on Uncertainty in Artificial Intelligence. 2006.
34. Jansen R., Yu H., Greenbaum D., Kluger Y., et al. A bayesian networks approach for predicting protein-protein interactions from genomic data. *Science*. 2003. vol. 302(5644). pp. 449–453.
35. Lacave C., Diez F.J. Knowledge Acquisition in PROSTANET – A Bayesian network for diagnosis prostate cancer. *Knowledge-Based Intelligent Information and Engineering Systems*. 2003. LNCS 2774. pp. 1345–1350.
36. Landuyt D., Lemmens P., D'hondt R., Broekx S., et al. An ecosystem service approach to support integrated pond management: A case study using Bayesian belief networks — Highlighting opportunities and risks. *Journal of Environmental Management*. 2014. vol. 145. pp. 79–87.

37. Léger A., Duval C., Farret R., Weber P., et al. Modeling of human and organizational impacts for system risk analyses. 9th International Probabilistic Safety Assessment and Management Conference. Hong Kong. 2008.
38. Lewis R.W., Ransing R.S. A semantically constrained Bayesian network for manufacturing diagnosis. *International Journal of Production Research*. 1997. vol. 35. no. 8. pp. 2171–2188.
39. Li P.C., Chen G.H., Dai L.C., Zhang L. A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRA frameworks. *Safety Science*. 2012. vol. 50(7). pp. 1569–1583.
40. Lucas P.J.F., De Bruijn N.C., Schurink K., Hoepelman I.M. A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. *ArtifIntell Med*. 2000. vol. 19(3). pp. 251–279.
41. Lucas P., van der Gaag L., Abu-Hanna A. Bayesian networks in biomedicine and healthcare. *Artificial Intelligence in Medicine*. 2004. vol. 30. pp. 201–214.
42. McCann R. K., Marcot B. G., Ellis R. Bayesian belief networks: applications in ecology and natural resource management. *Canadian Journal of Forest Research*. 2006. vol. 36. no. 12. pp. 3053–3062.
43. Mengshoel O.J., Darwiche A., Cascio K., Chavira M., et al. Diagnosing faults in electrical power systems of spacecraft and aircraft. Proceedings of IAAI-08. Chicago. 2008. pp. 1699–1705.
44. *Molloy B., Balfe N., Lowe E.* Developing Bayesian Belief Networks to Support Risk-Based Decision Making in Railway Operations. Conference Proceedings: Applied Human Factors and Ergonomics. Poland. 2014.
45. Nikovsky D. Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics. *IEEE Transactions on Knowledge and Data Engineering*. 2000. vol. 12. no. 4. pp. 509–516.
46. Oteniya L., Cowie J., Coles R. Diagnosis of Dementia and its Pathologies Using Bayesian Belief Networks. Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration (ICEIS 2006). Cyprus. 2006.
47. Radlinski L. A survey of bayesian net models for software development effort prediction. *International Journal of Software Engineering and Computing*. 2010. vol. 2. no. 2. pp. 95–109.
48. Ramoni M., Sebastiani P., Cohen P. Bayesian clustering by dynamics. *Mach Learn.* 2002. vol. 47. pp. 91–121.
49. Rao R.P.N. Neural Models of Bayesian Belief Propagation. The Bayesian Brain: Probabilistic Approaches to Neural Coding. Cambridge. MIT Press. 2007. 239 p.
50. Ritthaler M., Luger G., Young R. Bayesian Belief Networks for Astronomical Object Recognition and Classification in CTI-II. Astronomical Data Analysis Software and Systems XVI.ASP Conference Series. 2007. vol. 376. pp. 413–416.
51. Romessis C., Mathioudakis K. Bayesian network approach for gas path fault diagnosis. *Journal of engineering for gas turbines and power*. 2006. vol. 128(1). pp. 64–72.
52. Spiegelhalter D.J., Dawid A.P., Lauritzen S.L., Cowell R.G. Bayesian Analysis in Expert Systems. *Statistical Science*. 1993. vol. 8. no. 3. pp. 219–247. Available at: http://www.jstor.org/stable/2245959 (accessed: 02.11.2015).
53. Straub D. Natural hazards risk assessment using Bayesian networks. *Safety and Reliability of Engineering Systems and Structures*. 2005. pp. 2535–2542.
54. Tang A., Nicholson A., Jin Y., Han J. Using Bayesian belief networks for change impact analysis in architecture design. *Journal of Systems and Software*. 2007. vol. 80. no. 1. pp. 127–148.

55. Troldborg M., Aalders I., Towers W., Hallett P.D., et al. Application of Bayesian Belief Networks to quantify and map areas at risk to soil threats: Using soil compaction as an example. *Soil and Tillage Research*. 2013. vol. 132. pp. 56–68.
56. Trucco P., Cagno E, Ruggeri F., Grande O. A Bayesian Belief Network modeling of organizational factors in risk analysis: a case study in maritime transportation. *Reliability Engineering and System Safety*. 2008. vol. 93. pp. 823–834.
57. Twardy C.R., Nicholson A.E., Korb K.B., McNeil J. Epidemiological data mining of cardiovascular Bayesian networks. *Electronic Journal of Health Informatics*. 2006. vol. 1(1). pp. 3.
58. Verduijn M., et al. Prognostic Bayesian networks I: rationale, learning procedure, and clinical use. *Journal of Biomedical Informatics*. 2007. vol. 40(6). pp. 609–618.
59. Walsh S., Whitney P. A Graphical Approach to Diagnosing the Validity of the Conditional Independence Assumptions of a Bayesian Network Given Data. *Journal of Computational and Graphical Statistics*. 2012. vol. 21. no. 4. pp. 961–978.
60. Wasyluk H., Onisko A., Druzdzel M.J. Support of diagnosis of liver disorders based on a causal Bayesian network model. *Medical Science Monitor*. 2001. vol. 7. no. 1. pp. 327–332.
61. Yongli Z., Limin H., Jinling L. Bayesian network-based approach for power system fault diagnosis. *IEEE Transactions on Power Delivery*. 2006. vol. 21. pp. 634–639.
62. Zhang R., Moyle S., McKeever S., Bivens A. Performance problem localization in selfhealing, service-oriented systems using Bayesian networks. In Proceedings of the 2007 ACM symposium on Applied computing (SAC '07). 2007.

Торопова Александра Витальевна — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН), аспирант математико-механического факультета, Санкт-Петербургский государственный университет (СПбГУ). Область научных интересов: байесовские сети доверия, социокомпьютинг. Число научных публикаций — 15. alexandra.toropova@gmail.com; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7(812)328-3337.

Toropova Aleksandra Vital'evna — junior researcher of theoretical and interdisciplinary computer science laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), Ph.D. student, Saint Petersburg State University (SPbSU). Research interests: Bayesian belief networks, social computing. The number of publications — 15. alexandra.toropova@gmail.com; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)328-3337.

Поддержка исследований. Работа выполнена при частичной финансовой поддержке РФФИ (проекты 14-01-00580-а, 15-01-09001-а).

Acknowledgements. This research is supported by RFBR (grants 14-01-00580-a, 15-01-09001-a).

РЕФЕРАТ

Торопова А.В. **Подходы к диагностике согласованности данных в байесовских сетях доверия.**

Байесовские сети доверия предоставляют возможность объединения нескольких видов информации, например полученной от экспертов или статистически, позволяют работать с неполной или неточной информацией, обладают наглядностью и другими полезными свойствами. Благодаря этому они стали популярным и весьма эффективным средством. Однако во многих областях исследования исходные данные, которые предоставляются экспертами, могут быть не согласованы, и поэтому в некоторых задачах следует использовать инструменты для проверки их согласованности.

Цель работы – систематизировать с помощью обзора примеры и задачи, в которых применяются байесовские сети доверия, чтобы оценить в какой степени в этих задачах учитывается диагностика согласованности исходных данных, и насколько важным является ее применение.

В работе рассмотрены примеры применения аппарата байесовских сетей доверия в различных областях.

Задачи медицины и здравоохранения подразделяются на четыре основных направления: диагностика, прогнозирование, лечение заболеваний и обнаружению функциональных взаимодействий между различными объектами. Диагностика согласованности исходных данных может потребоваться в первых трех, так как значительная их часть предоставляется экспертами, которые могут быть субъективны или ошибаться.

В задачах экологии в основном используется смесь экспертных данных и точных технических показаний, инструменты диагностирования экспертных данных могут помочь достичь более точных результатов.

В области функциональной безопасности систем исходные данные обычно организованы таким образом, что диагностика их согласованности не требуется.

В экономики и риск-анализе частым является интеграция различной информации, в частности статистических и полученных от экспертов данных. Данные полученные от экспертов могут содержать противоречия, поэтому следует иметь инструмент, позволяющий определить насколько согласованы такие данные.

В области социологических исследований так как часто исследования проводятся на основе опросов различных групп населения, данные предоставляемые ими могут оказаться несогласованными и требуют проверки. В задачах анализа социальных сетей, где учитываются связи между членами сетей, диагностика согласованности данных не требуется.

Таким образом, в работе показана необходимость создания и использования инструментов диагностики согласованности исходных данных в некоторых областях исследований.

SUMMARY

Toropova A.V. **Approaches to the Data Coherence Diagnosis in Bayesian Belief Network Models.**

Bayesian belief networks provide the ability to combine different types of information, e.g. statistical or expert data, allow working with incomplete or inaccurate information; they have clarity and other useful properties. Due to this, Bayesian belief networks have become a popular and highly effective tool in many fields of research. However, in many research areas data provided by the experts can be incoherent, and so in some tasks one should use tools to verify their coherence.

The purpose of this work is to systematize problems and examples using Bayesian belief networks by reviewing and to assess their use of data coherence diagnosis and its importance. The paper discusses examples of application of the Bayesian belief networks in different research areas.

Health-care and medicine problems are divided into four base areas: diagnosis, prognosis, treatment selection and discovering functional interactions between different objects. Initial data coherence diagnosis can be required in the first three, as much of this data is provided by experts who can be subjective or wrong.

The ecology problems contain a mix of an expert data and precise technical indications; expert data coherence diagnostic tools can help achieve more accurate results.

In the functional safety systems, data is mainly organized in such a way that the coherence diagnosis is not required.

In economics and risk analysis, it is common to integrate different kinds of information, in particular data that is statistical or received from the experts. The data obtained from the experts may contain contradictions, so tools for discovering incoherent data should be used.

In sociology, researchers usually retrieve initial data from the different population groups' surveys, so they can provide incoherent or even intentionally wrong data. Hence, these data is requiring a coherence diagnosis. The SNA problems consider links between network members, therefore coherence diagnosis is not necessary.

Thus, the work shows the necessity to create and use the coherence diagnostic tools of input data in different research areas.

УДК 004.056

И.В. ГАВРИЛОВ

МЕТОДИКА ОЦЕНИВАНИЯ КАЧЕСТВА МАСКИРУЮЩЕГО ШУМА

Гаврилов И.В. **Методика оценивания качества маскирующего шума.**
Аннотация. Эффективная защита конфиденциальной речевой информации генераторами маскирующего шума является достаточно важной задачей для большинства государственных и коммерческих учреждений. Тем не менее, в настоящее время нет единого подхода к оценке качества маскирующих шумов для зашумления речевой информации, а существующие методики нуждаются в серьёзной доработке. В статье представлена модифицированная методика оценивания качества шума, используемого при маскировании речевой информации. Данная методика на основе введённого параметра равномерности амплитудного спектра позволяет учитывать и рассчитывать степень провалов в частотной области шумового сигнала.
Ключевые слова: маскирующий шум, энтропия, частотный спектр сигнала, средства активной защиты.

Gavrilov I.V. **Method of Evaluating the Quality of Masking Noise.**
Abstract. Effective protection of confidential information by using masking noise generators is quite an important task for most government and commercial institutions. However, there is currently no common approach to the assessment of the quality of voice information noise masking and the existing techniques require major improvements. The article presents a modified method of evaluating the quality of noise used for masking voice information. This methodology based on the entered uniformity parameter of amplitude spectrum allows one to consider and calculate the degree of failure in the frequency domain of the noise signal.
Keywords: masking noise, entropy, frequency spectrum of signal, means of active protection.

1. Введение. В современном мире с нарастающими объёмами обрабатываемых данных растёт и количество речевой информации в государственных учреждениях и на предприятиях, в процессе проведения различных совещаний, конференций, собраний, пленумов, заседаний и при ведении переговоров. Зачастую обсуждаются сведения конфиденциального характера, которые могут быть отнесены к коммерческой или государственной тайне. Очевидно, что в таких условиях необходимо обеспечивать гарантированную защиту упомянутых сведений, которую можно организовать с использованием активных средств, например, таких, как генераторы маскирующего шума, описанные в [1–3]. Но зашумлённый информативный сигнал может быть подвергнут фильтрации и в случае некачественного маскирования злоумышленник получит доступ к защищаемым сведениям. Поэтому возникает важная задача, связанная с оценкой качества шумового сигнала, порождаемого средствами активной защиты.

2. Критерии качества шумового сигнала для защиты информации. Оценивание защищённости речевой информации в настоящее время проводится в соответствии с методиками, изложенными

в [4–6], путём расчёта словесной или формантной разборчивости речи, по которым косвенно можно судить о качестве маскирующего шума.

Для определения оценочных характеристик маскирующего шума используются информационные [7] и энергетические [8, 9] критерии. Первая группа критериев рассматривает статистические параметры шумовых сигналов во временной области и позволяет непосредственно определить числовой коэффициент качества шума. На основе расчёта математического ожидания, дисперсии и энтропии мгновенных значений временных отсчётов и их огибающей вычисляется степень приближения к некоторым эталонным распределениям. Такие методы направлены на нахождение степени неопределённости мгновенных значений шумовых сигналов, выражаемых, например, через энтропийный коэффициент качества маскирующего шума [7].

Критерии из второй группы для гарантированной защиты информации используют постулат о необходимости превышения энергетики шума над маскируемым сигналом. Поэтому с целью проверки качества шума используются интегральные показатели, учитывающие превышение уровня шума над уровнем информативного сигнала. Например, весь частотный диапазон маскирующего шума может разбиваться на несколько октавных полос, на средних частотах каждой из которых измеряется уровень шума [8].

С точки зрения энергетической эффективности генерации маскирующих шумов, а также для непосредственного определения их вероятностных свойств наибольший интерес представляют информационные критерии. Поэтому в статье рассмотрена методика оценки качества маскирующего шума, подробно описанная в [7], которая относится к группе информационных критериев.

3. Методика оценки качества маскирующего шума. Способы оценки качества маскирующего шума различного рода были описаны в следующих источниках [7, 10–13]. Данные способы сводятся к ряду вычислительных операций, производимых с квантованными измеренными значениями электрического сигнала, к которому преобразуется маскирующий шум. Основу данных способов составляет расчёт меры неопределённости (энтропии) закона распределения мгновенных значений маскирующего шума, а также энтропии закона распределения значений огибающей шумового сигнала.

В данных работах вводится понятие энтропийного коэффициента качества шума. Указанные коэффициенты рассчитываются относительно некоторых эталонных законов распределения. Для мгновенных значений маскирующего шума в условиях ограничений, накладываемых на среднюю мощность, эталонным является нормальный закон

распределения, а для огибающей нормально распределённых мгновенных значений маскирующего шума – закон распределения Релея [7].

Данную методику можно представить в виде последовательности приведённых далее основных операций.

1. Вычисление математического ожидания и среднеквадратического значения напряжения шумового сигнала по полученным в результате процедур дискретизации и квантования мгновенным значениям напряжения и рассчитанным вероятностям пересечения уровней квантования.

2. Определение энтропийного коэффициента качества маскирующего шума по мгновенным значениям относительно параметров нормального распределения по следующей формуле:

$$\eta^M = \frac{e^H}{\sqrt{2\pi\sigma^2}}, \qquad (1)$$

где σ – среднеквадратическое значение напряжения шумового сигнала; H – энтропия закона распределения мгновенных значений напряжения шумового сигнала.

3. Определение энтропийного коэффициента на основе вероятностей распределения значений огибающей по уровням квантования, математического ожидания натурального логарифма значений огибающей и второго момента закона распределения значений огибающей маскирующего шума.

$$\eta^0 = \frac{e^{H^0}}{e^{H^P}} = e^{H^0 + m - 2\ln r - \frac{\sigma_0}{2r^2}}, \qquad (2)$$

где H^0 – энтропия закона распределения огибающей напряжения шумового сигнала; m – математическое ожидание натурального логарифма значений напряжения огибающей; r – параметр закона распределения Релея; σ_0 – второй момент закона распределения значений напряжения огибающей шумового сигнала.

4. Расчёт общего энтропийного коэффициента качества маскирующего шума производится перемножением полученных с помощью выражений (1) и (2) энтропийных коэффициентов.

В результате расчёта относительного энтропийного коэффициента качества без затраты значительных ресурсов появляется возможность оценивания маскирующих шумов.

Основным недостатком описанного выше подхода является учёт только временных составляющих маскирующего шума. Данный факт

не позволяет проводить анализ неравномерности частотного спектра, поэтому возможно наличие провалов и подъёмов шумовых сигналов в различных частотных областях, что может серьёзно повлиять на качество маскировки сигналов [2, 14, 15].

4. Уточнение энтропийного коэффициента качества маскирующего шума. Для устранения указанного недостатка предлагается доработать методику оценки маскирующего шума, учитывая его частотные свойства.

В соответствии с требованиями нормативных документов по технической защите информации спектр маскирующего шума должен соответствовать спектру маскируемого сигнала [14], другими словами, наиболее эффективным маскирующим шумом является формантоподобный [6].

Речевой сигнал представляет из себя шумоподобный процесс со случайной модуляцией. Усреднённые характеристики типовых речевых сигналов экспериментально получены. Средние спектральные уровни и артикуляционные параметры для равноартикуляционных и октавных полос речи представлены в [5]. С практической точки зрения зашумлённость речевого сигнала легче оценивать в октавных полосах, при этом погрешность измерений по сравнению с равноартикуляционными полосами составит менее 10%.

Примеры спектров речи дикторов (женщины и мужчины) с указанием семи октавных полос представлены на рисунках 1 и 2.

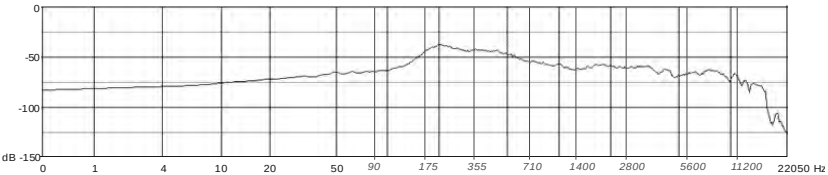


Рис. 1. Усреднённый спектр речи диктора (женщины)

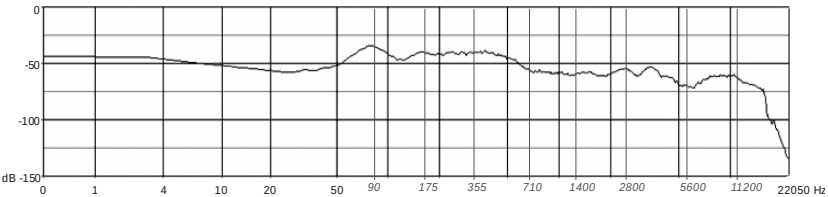


Рис. 2. Усреднённый спектр речи диктора (мужчины)

Рассматривая особенности работы слухового аппарата человека, необходимо упомянуть критические полосы слуха, соответствующие разрешающей способности уха по частоте [16]. Шириной критических

полос определяется слуховая чувствительность уха. Критические полосы для частот до 1 кГц в среднем составляют 35-50 Гц. В то же время минимальная ширина октавной полосы составляет 125 Гц [5]. Это означает, что критическая полоса в несколько раз меньше октавной, поэтому возникает важная задача по оценке равномерности маскирующего шума в отдельных октавных полосах.

Также из отмеченного выше следует, что наиболее приемлемым маскирующим шумом будет случайный шум с нормальным распределением плотности вероятности мгновенных значений и равномерным частотным спектром в октавных полосах. [2, 14, 15].

Для определения степени равномерности спектра маскирующего шума в отдельных октавных полосах был введён коэффициент равномерности амплитудного спектра маскирующего шума [17]. Указанный коэффициент оценки маскирующего шума для выбранного частотного диапазона рассчитывается на основе усреднения относительных коэффициентов схождения к среднему значению.

Подробное описание процесса определения коэффициента равномерности в октаве в виде последовательности действий можно продемонстрировать следующим образом.

1. Нахождение среднего значения амплитуд спектральных составляющих маскирующего шума, например, с помощью быстрого преобразования Фурье.

2. Разбиение требуемого частотного диапазона на *M* частотных областей таким образом, чтобы в каждой области оставалось более двух значений амплитуд спектральных составляющих (рисунок 3).

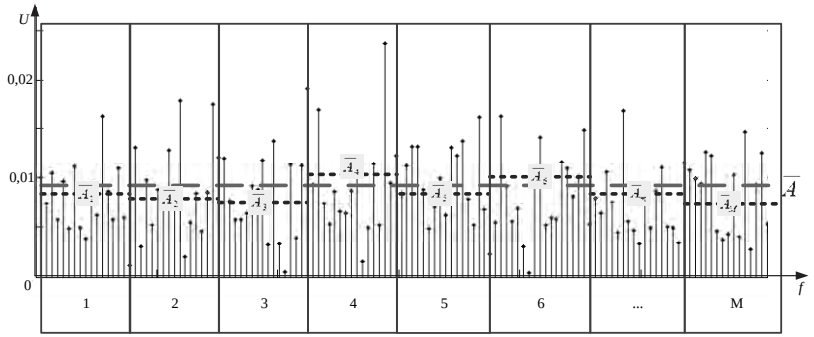


Рис. 3. Иллюстрация сущности процесса определения коэффициента равномерности

3. Расчёт относительных коэффициентов схождения к среднему значению амплитуд спектральных составляющих $\overline{A}$ для каждой m-ой частотной области по следующей формуле:

$$K_m = \frac{\left| \overline{A} - \frac{1}{p_m} \sum_{j=1}^{p_m} A_j \right|}{\overline{A}}, \quad (3)$$

где $\overline{A} = \frac{1}{n} \sum_{i=1}^{n} A_i$ – среднее значение амплитуд; A_j – значение амплитуды спектральных составляющих; p_m – количество спектральных составляющих в границах m-ой частотной области; n – количество спектральных составляющих в заданной ограниченной полосе частот.

4. Вычисление коэффициента равномерности амплитудного спектра через усреднённое значение относительных коэффициентов схождения к среднему значению в соответствии со следующим выражением:

$$\overline{K} = 1 - \frac{1}{M} \sum_{m=1}^{M} K_m . \quad (4)$$

С использованием полученного коэффициента равномерности амплитудного спектра корректируется общий энтропийный коэффициент качества шума.

5. Оценка эффективности предложенной методики определения качества маскирующего шума. Целью модифицированной методики является повышение точности оценивания маскирующего шума. Для расчёта получаемого эффекта предлагается использовать значение коэффициента отклонения от среднего значения. Указанный коэффициент демонстрирует отклонение спектра маскирующего шума от равномерно распределённого спектра в октавных полосах с энергетикой, эквивалентной рассматриваемому маскирующему шуму [17].

$$\overline{K_0} = \frac{1}{M} \sum_{m=1}^{M} \frac{\left| \frac{1}{n} \sum_{i=1}^{n} A_i - \frac{1}{p_m} \sum_{j=1}^{p_m} A_j \right|}{\frac{1}{n} \sum_{i=1}^{n} A_i} \cdot 100\%, \quad (5)$$

где M – количество частотных областей; n – количество спектральных составляющих в заданной ограниченной полосе частот; p_m – количест-

во спектральных составляющих в границах *m*-ой частотной области; *A* – значение амплитуды спектральных составляющих.

На примере отрезка записи маскирующего шума, амплитудно-частотный спектр которого представлен на рисунке 4, произведён расчёт коэффициента равномерности.

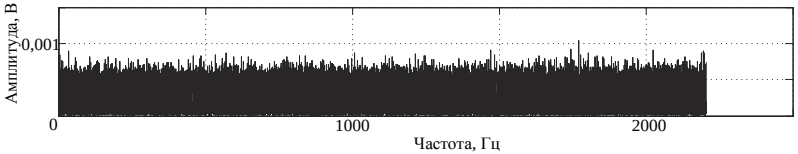


Рис. 4. Амплитудно-частотный спектр экспериментального маскирующего шума

При вычислении общего энтропийного коэффициента качества, порядок действий для определения которого представлен в п.3, по выражениям (1) и (2) получено значение 0,1.

Дальнейший порядок действий соответствует модифицированной методике оценивания качества маскирующего шума, рассмотренной в п.4. Спектр указанного маскирующего шума представлен 262 145 спектральными составляющими. Весь частотный диапазон в данном случае был разделён на 20000 частотных областей. Иллюстрацией к производимым расчётам для первых 15 частотных областей является рисунок 5.

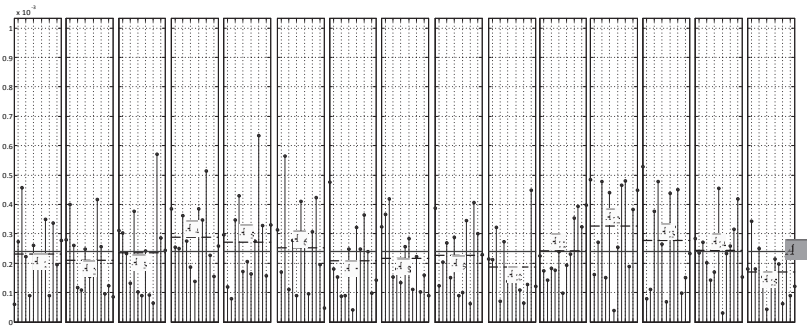


Рис. 5. Первые частотные области амплитудно-частотного спектра экспериментального маскирующего шума

Для каждой частотной области рассчитаны относительные коэффициенты схождения к среднему значению по формуле (3). В соответствии с выражением (4) для коэффициента равномерности получено значение 0,877. В результате корректировки значение энтро-

пийного коэффициента качества для рассмотренного маскирующего шума составило 0,087.

По выражению (5) эффект в результате применения данного способа оценки при корректировке энтропийного коэффициента рассматриваемого маскирующего шума составил 12,3%. Таким образом была повышена точность оценивания маскирующего шума, что свидетельствует о достижении поставленной цели.

6. Заключение. В результате проведённых исследований разработана модифицированная методика оценивания качества маскирующего шума, которая позволяет уточнить существовавший подход к определению качества маскирующего шума [7], создаваемых средствами активной защиты речевой информации, через расчёт энтропийного коэффициента качества этого шума.

Таким образом, в результате проведённых исследований получена возможность вычисления оценочного показателя, учитывающего качество маскирующего шума не только во временной, но и в частотной области.

Данные исследования позволят в будущем более объективно оценивать качество маскирующего шума, а на основе полученных оценок делать вывод о функционировании средств активной защиты.

Литература

1. *Зайцев А.П., Шелупанов А.А., Мещеряков Р.В. и др*. Технические средства и методы защиты информации: учеб. для вузов // М.: ООО «Издательство Машиностроение». 2009. 508 с.
2. *Халяпин Д. Б.* Защита информации. Вас подслушивают? Защищайтесь! // М.: НОУ ШО «Баярд». 2004. 432 с.
3. *Хорев А. А.* Защита информации от утечки по техническим каналам. Часть1. Технические каналы утечки информации : учеб. пособие // М. : Гостехкомиссия России. 1998. 320 с.
4. *Покровский Н.Б.* Расчёт и измерение разборчивости речи // М.: "Связьиздат", 1962. 392 с.
5. *Железняк В.К., Макаров Ю.К., Хорев А.А.* Некоторые методические подходы к оценке эффективности защиты речевой информации // Специальная техника. 2000. № 4. С. 39–45.
6. *Иванов А.В., Рева И.Л., Трушин В.А.* Реализация оптимальной помехи при защите речевой информации от утечки по акустическому и виброакустическому каналам // Научный вестник НГТУ. 2011. №4(45). С. 151–154.
7. *Герасименко В.Г., Лаврухин Ю.Н., Тупота В.И.* Методы защиты акустической речевой информации от утечки по техническим каналам // М. : РЦИБ "Факел". 2008. 258 с.
8. *Дворянкин С.В., Макаров Ю.К., Хорев А.А.* Обоснование критериев эффективности защиты речевой информации от утечки по техническим каналам // Защита информации. Инсайд. 2007. № 2. С. 18–25.
9. *Козлачков С.Б.* Дополнительные критерии оценки защищённости речевой информации // Спецтехника и связь. 2011. № 2. С. 44–47.

10. *Тупота В.И. и др.* Способ оценки качества маскирующих частотно-модулированных шумовых помех // Патент РФ № 2346390. 2009. Бюл. № 4. 6 с.
11. *Тупота В.И. и др.* Способ оценки качества маскирующего акустического (виброакустического) шума // Патент РФ № 2350023. 2009. Бюл. № 8. 10 с.
12. *Тупота В.И. и др.* Способ оценки качества маскирующих амплитудно-модулированных шумовых помех // Патент РФ № 2351076. 2009. Бюл. № 9. 6 с.
13. *Тупота В.И. и др.* Способ оценки качества маскирующих прямошумовых помех // Патент РФ № 2353057.2009. Бюл. № 11. 9 с.
14. *Бузов Г.А., Калинин С.В., Кондратьев А.В.* Защита от утечки информации по техническим каналам: учеб. пособие // М.: Горячая линия – Телеком. 2005. 416 с.
15. *Каторин Ю.Ф., Куренков Е.В., Лысов А.В., Остапенко А.Н.* Энциклопедия промышленного шпионажа // СПб: ООО «Издательство Полигон». 2000. 512 с.
16. *Сапожков М.А.* Электроакустика: учебник для вузов. М.: «Связь». 1978. 272 с.
17. *Гаврилов И.В. и др.* Способ оценки качества маскирующего шума // Патент РФ № 2550353. 2015. Бюл. № 13. 13 с.

References

1. Zaicev A.P., Shelupanov A.A., Mescheryakov R.V. i dr. *Tehnicheskie sredstva i metody zaschity informacii: uchebnik dlya vuzov* [Hardware and methods of information security: textbook for high schools]. M.: OOO «Izdatel'stvo Mashinostroenie». 2009. 508 p. (In Russ.).
2. Halyapin D. B. *Zaschita informacii. Vas podslushivayut? Zaschischaites'!* [Information security. Did you overhear? Defend yourself!]. M.: NOU ShO «Bayard». 2004. 432 p.
3. Horev A. A. *Zaschita informacii ot utechki po tehnicheskim kanalam. Chast'1. Tehnicheskie kanaly utechki informacii*: *ucheb. posobie* [Information protection against leakage via technical channels. Part 1. Technical channels of information leakage: tutorial]. M.: Gostehkomissiya Rossii. 1998. 320 p. (In Russ.).
4. Pokrovskiy N.B. *Raschyot i izmerenie razborchivosti rechi* [Calculation and measurement of speech intelligibility]. M.: «Sviaz`izdat». 1962. 392 p. (In Russ.).
5. Zhelezniak V.K., Makarov U.K., Horev A.A. [Some methodological approaches to evaluating the effectiveness of the protection of the speech information]. *Spetcial`naia tekhnika – Special equipment.* 2000. vol 4. pp. 39–45. (In Russ.).
6. Ivanov A.V., Reva I.L., Trushin V.A. [Implementation of optimal interference protection of speech information leakage on acoustic and vibro-acoustic channels]. *Nauchny`i` vestneyk NGTU – Science bulletin of NSTU*. 2011. vol. 4(45). pp. 151–154. (In Russ.).
7. Gerasimenko V.G., Lavruhin Yu.N., Tupota V.I. Metody zaschity akusticheskoi rechevoi informacii ot utechki po tehnicheskim kanalam [Methods of protection of the acoustic speech information leakage via technical channels]. M.: RCIB "Fakel". 2008. 258 p. (In Russ.).
8. Dvoryankin S.V., Makarov Yu.K., Horev A.A. [Justification of the criteria the protection of the speech information leakage via technical channels]. *Zaschita informacii. Insaid – Information security Inside.* 2007. vol. 2. pp. 18–25. (In Russ.).
9. Kozlachkov S.B. Dopolnitel'nye kriterii ocenki zaschischennosti rechevoi informacii [Additional criteria for evaluating the security of voice data]. *Spectehnika i svyaz' – Special equipment and communication.* 2011. vol. 2. pp. 44–47. (In Russ.).
10. Tupota V.I. et al. *Sposob ocenki kachestva maskirujushhih chastotno-modulirovannyh shumovyh pomeh* [Method for assessment of quality of masking frequency-modulated noise jamming]. Patent RF. no. 2346390. 2009. Bull. no. 4. 6 p. (In Russ.).

11. Tupota V.I. et al. *Sposob ocenki kachestva maskirujushhego akusticheskogo (vibroakusticheskogo) shuma* [Masking acoustic (vibroacoustic) noise quality test]. Patent RF. no. 2350023. 2009. Bull. no. 8. 10 p. (In Russ.).
12. Tupota V.I. et al. *Sposob ocenki kachestva maskirujushhih amplitudno-modulirovannyh shumovyh pomeh* [Method of evaluating quality of masking amplitude-modulated noise interference]. Patent RF. no. 2351076. 2009. Bull. no. 9. 6 p. (In Russ.).
13. Tupota V.I. et al. *Sposob ocenki kachestva maskirujushhih prjamoshumovyh pomeh* [Assessment method of masking direct noise interference quality]. Patent RF. no. 2353057. 2009. Bull. no. 11. 9 p. (In Russ.).
14. Buzov G.A., Kalinin S.V., Kondrat'ev A.V. *Zaschita ot utechki informacii po tehnicheskim kanalam: ucheb. posobie* [Protection against leakage of information through technical channels: tutorial]. M.: Goryachaya liniya – Telekom. 2005. 416 p. (In Russ.).
15. Katorin Yu.F., Kurenkov E.V., Lysov A.V., Ostapenko A.N. *Enciklopediya promyshlennogo shpionaja. Pod obsch. red. Kurenkova E.V.* [Encyclopaedia of industrial espionage. edited by Kurenkov E.V.]. SPb: OOO «Izdatel'stvo Poligon». 2000. 512 p. (In Russ.).
16. Sapozhkov M.A. *E`lektroakustika: uchebnik dlia vuzov* [Electroacoustics: textbook for high schools.]. M.: «Sviaz`». 1978. 272 p. (In Russ.).
17. Gavrilov I.V. et al. *Sposob ocenki kachestva maskirujushhego shuma* [Quality assessment method of masking noise]. Patent RF. no. 2550353. 2015. Bull. № 13. 13 p. (In Russ.).

Гаврилов Илья Вячеславович — аспирант, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: системы активной защиты информации. Число научных публикаций — 5. ilya_vch@pisem.net; Приборостроительная, 35, Орел, 302034; р.т.: +7(4862)549533.

Gavrilov Ilya Vyacheslavovich — Ph.D. student, The Academy of Federal Security Guard Service of the Russian Federation. Research interests: system of active information security. The number of publications — 5. ilya_vch@pisem.net; 35, Priborostroitelnaya Street, Orel, 302034, Russia; office phone: +7(4862)549533.

РЕФЕРАТ

Гаврилов И.В. **Методика оценивания качества маскирующего шума.**

В результате увеличения количества обрабатываемой речевой информации конфиденциального характера повышаются требования к её защите. А в качестве активных средств защиты используются генераторы маскирующего шума. Поэтому существует важная задача, связанная с оценкой качества шума.

После проведённого анализа методов определения качества шумового сигнала установлено, что одним из наиболее перспективных является метод, основанный на определении энтропийного коэффициента качества маскирующего шума. Но в данном методе есть существенный недостаток: он не позволяет оценивать частотную составляющую маскирующего шума.

Поэтому в работе предложен подход к определению частотных свойств маскирующего шума. Установлено, что наилучшими шумовыми свойствами обладает маскирующая помеха с равномерным частотным спектром.

В статье предлагается для оценки качества частотной составляющей маскирующего шума использовать коэффициент равномерности амплитудного спектра, который вычисляется через рассчитываемое усреднённое значение относительных коэффициентов схождения к среднему. В свою очередь энтропийный коэффициент качества для временной составляющей корректируется с учётом значения коэффициента равномерности для частотной составляющей.

Предложенный подход к определению качества маскирующего шума за счёт дополнительного учёта огибающей частотного спектра позволяет повысить точность проводимой оценки. Это показано в работе на примере расчёта коэффициента равномерности и корректировки энтропийного коэффициента для некоторой реализации маскирующего шума.

SUMMARY

Gavrilov I.V. **Method of Evaluating the Quality of Masking Noise.**

Increase in the amount of processed confidential voice information has led to increased requirements for its protection. Noise generators are used as active means of protection. Therefore, an important task arising is that of the assessment of the noise quality.

After conducted analysis of methods of determining the quality of a noise signal, it is established that one of the most promising methods is a method based on the determination of the entropy coefficient of the masking noise quality. However, this method has a major drawback: it does not evaluate the masking noise frequency component.

Therefore, in this paper we propose an approach to the determination of the frequency properties of masking noise. Masking interference with uniform frequency spectrum was found to have the best noise characteristics.

The paper proposes to assess the quality of the masking noise frequency component by using the uniformity coefficient of amplitude spectrum, which is calculated by the estimated average values of the relative coefficients regression to the mean. In its turn, the entropy quality coefficient for the time component is adjusted taking into account the values of the uniformity coefficient for the frequency component.

The proposed approach to the determination of the quality of masking noise due to additional accounting of an envelope of the frequency spectrum allows improving the accuracy of the assessment. In the paper, it is shown using the example of calculating the uniformity coefficient and adjusting the entropy coefficient for the implementation of masking noise.

УДК 004.934: 004.42: 61:00.7

Д.А. Вольф, Р.В. Мещеряков

МОДЕЛЬ И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ СИНГУЛЯРНОГО ОЦЕНИВАНИЯ ЧАСТОТЫ ОСНОВНОГО ТОНА РЕЧЕВОГО СИГНАЛА

Вольф Д.А., Мещеряков Р.В. **Модель и программная реализация сингулярного оценивания частоты основного тона речевого сигнала.**

Аннотация. В статье рассматривается сингулярная модель оценивания частоты основного тона речевого сигнала, а также ее программная реализация. Применение модели сингулярного оценивания частоты основного тона позволяет уменьшить вычислительную сложность алгоритмов анализа речевого сигнала путем аппроксимации края сингулярного спектра и обеспечить меньшее количество ошибок оценивания частоты основного тона за счет использования сингулярной модели вокализированного сегмента речи, учитывающей нестационарные параметры основного тона с помощью собственных чисел. Программная реализация модели используется в модуле расчетов комплекса программ речевой реабилитации онкологических больных после резекции гортани.

Ключевые слова: оценивание частоты основного тона речевого сигнала, сингулярный спектральный анализ речи, модель, программная реализация.

Volf D.A., Meshcheryakov R. V. **Software Implementation of a Singular Meter of the Pitch Frequency of a Speech Signal.**

Abstract. The article deals with software implementation of the evaluation of the pitch frequency of the speech signal based on the mathematical apparatus for singular spectral analysis. The program is used in calculation module of a program complex for speech rehabilitation of cancer patients after resection of larynx used in rehabilitation training of patients after complete or partial loss of sounding speech as a result of laryngectomy.

Keywords: estimation of the pitch frequency of the speech signal; singular spectrum analysis of speech; model; software implementation.

1. Введение. Поставленная в работе задача определения частоты основного тона речевого сигнала, включая распределение амплитуд, периодов и начальных фаз гармоник, образующих сложный полигармонический сигнал, остается все еще нерешенной и активно исследуемой в области речевых технологий [1–3]. Существующие алгоритмы оценивания ЧОТ [4–7] позволяют проводить анализ статистических данных без учета особенностей речеобразования и речевосприятия, связанных с анатомией и физиологией человека, так как методы анализа [8], лежащие в их основе, ограничены периодической (стационарной) моделью речевого сигнала, которая подразумевает точное повторение периода и амплитуды основного тона и не допускает их изменения на протяжении окна анализа. В свою очередь, это влияет на точность результатов оценивания ЧОТ [9].

Исходя из данной проблемы, появляется мотивация к разработке такой модели, которая позволит осуществлять учет нестационарных

амплитуд, периодов и фаз гармоник, входящих в речевой сигнал. С другой стороны, повышение точности вычисления ЧОТ приводит к увеличению вычислительной сложности [10]. Таким образом, разработка новых методов анализа речи для задач оценивания частоты основного тона речи, является актуальной [11].

Целью настоящей работы является получение модели оценивания частоты основного тона речевого сигнала при оптимальной временной обработке с учетом особенностей речеобразования и речевосприятия, связанных с анатомией и физиологией человека, а также получения ее программной реализации. Новизна данной работы заключается в применении математического аппарата сингулярного спектрального анализа к обработке речевых сигналов.

В рамках базовой части государственного задания ТУСУР (проект № 3657 от 2015г.) для НИИ Онкологии г. Томска разработан программный комплекс речевой реабилитации онкологических больных после резекции гортани (Свидетельство о государственной регистрации программы для ЭВМ № 2015618857 – "Программа речевой реабилитации больных после резекции гортани"). Разработанный программный комплекс состоит из семи модулей каждый из которых представляет собой черный ящик, который принимает на вход речевые данные, обрабатывает их и возвращает обратно интерпретируемый результат. Одним из ключевых модулей является модуль расчетов, в котором решаются задачи вычисления параметров речи. Одной из решаемых задач является оценивание частоты основного тона (ЧОТ) речевого сигнала.

2. Сингулярная модель вокализированного сегмента речи и сингулярная модель оценивания частоты основного тона. Особенность предлагаемого метода оценивания ЧОТ заключается в разложение речевого сигнала в элементарный спектр временных рядов (квазигармоник) посредством сингулярного спектрального анализа с последующим выбором квазигармонической составляющей, соответствующей основному тону речи. Рассмотрим прямую задачу. Пусть ряд S_N, полученный в результате процедуры дискретизации речевого сигнала $S(t)$, принимается в качестве фонемного ряда. С фонемным рядом проводится процедура Ганкелизации [12, 13]:

$$\mathbf{A}=[S_{i-1},...,S_{i+L-1}]^{\mathrm{T}},\ 1\le i\le K,\ K = N - L + 1, \qquad (1)$$

таким образом получается траекторная матрица **A**, состоящая из K векторов вложений длины L. Для траекторной матрицы (1) вычисляется матричное разложение вида:

$$\mathbf{A} = \sum_{i=0}^{L-1} \mathbf{A}^{<i>} = \sum_{i=0}^{L-1} (\sqrt{\lambda_i}\,\mathbf{u}^{<i>})[\mathbf{x}^{<i>}]^{\mathrm{T}}, \qquad (2)$$

где: λ_i – i-е собственное значение ковариационной матрицы $\mathbf{A}\mathbf{A}^{\mathrm{T}}$;
$\mathbf{u}^{<i>}$ – i-й собственный вектор ковариационной матрицы $\mathbf{A}\mathbf{A}^{\mathrm{T}}$;
$\mathbf{x}^{<i>}$ – i-й собственный вектор (главных компонент), образованный строками матрицы **A**.

Для произведения векторов в (2) вычисляется матричное усреднение по диагонали:

$$\mathbf{T}_j^{<n>} = \begin{cases} \dfrac{1}{j+1} \displaystyle\sum_{i=0}^{j} [\sqrt{\lambda_n}\,\mathbf{u}^{<n>}\mathbf{x}^{<n>\mathrm{T}}]_{iK+j-i}, 0 \le j < L; \\ \dfrac{1}{L} \displaystyle\sum_{i=0}^{L-1} [\sqrt{\lambda_n}\,\mathbf{u}^{<n>}\mathbf{x}^{<n>\mathrm{T}}]_{iK+j-i}, L \le j < K; \\ \dfrac{1}{N-j} \displaystyle\sum_{i=0}^{L-1-(j-K)} [\sqrt{\lambda_n}\,\mathbf{u}^{<n>}\mathbf{x}^{<n>\mathrm{T}}]_{(j-K+i)K+K-1-i}, K \le j < N. \end{cases} \qquad (3)$$

в результате которого образуется матрица временных рядов $\mathbf{T}_{L,N}$ в строках которого содержится квазигармонический спектр. Аналогично тому, как в гармонических моделях осуществляется проекция речевого сигнала в гармонический базис (например, в преобразованиях Фурье) [14, 15], так и в прямой задаче сингулярного спектрального анализа речи осуществляется проекция в базис собственных векторов. Рассмотрим обратную задачу. Пусть имеется квазигармонический спектр (3), тогда сумма j-х квазигармоник данного спектра будет равна исходному фонемному ряду:

$$S_N = \sum_{n=0}^{L-1} \mathbf{T}_j^{<n>}, \; j = 0,\ldots,N-1. \qquad (4)$$

Пусть для некоторой последовательности i=0,1,… собственные числа λ_i, $\mathbf{u}^{<i>}$, $\mathbf{x}^{<i>}$ – эмпирически найденные величины, образуют совокупность параметров для образования звуков речи, тогда для произведения:

$$\mathbf{A}_i = \sqrt{\lambda_i}\,\mathbf{u}^{<i>}[\mathbf{x}^{<i>}]^{\mathrm{T}}, \; i=0,\ldots,$$

выражение (3) можно принять в качестве синтезатора акустических сигналов, генерируемых речеобразующим трактом (рисунок 1). Без решения прямой задачи синтезирование параметров λ_i, $\mathbf{u}^{<i>}$, $\mathbf{x}^{<i>}$ в качестве резонаторов речеобразующего тракта является достаточно сложным процессом [16, 17]. Тем не менее, систему:

$$\begin{cases}(3);\\(4).\end{cases} \qquad (5)$$

можно принять в качестве сингулярной модели вокализированного сегмента речевого сигнала для решения задачи оценивания ЧОТ.

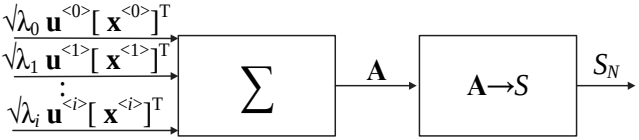


Рис. 1. Модель сингулярного синтезатора речи

Таким образом, можно сформулировать следующие фундаментальные тезисы для сингулярной модели вокализированной речи:

1. Система (5) наглядным образом показывает, что принимаемая сингулярная модель вокализированного сегмента речевого сигнала позволяет анализировать (рассматривать) речевой сигнал, в котором неизвестны амплитуды, периоды и начальные фазы всех гармоник.

2. Если речеобразующий тракт рассматривать как систему акустических резонаторов, тогда каждая *i*-я тройка чисел (λ_i, $\mathbf{u}^{<i>}$, $\mathbf{x}^{<i>}$), как отдельный параметр *i*-го резонатора, содержит информацию об индивидуальном акустическом различие, так как пространство собственных векторов **x** образует нестационарных базис, в который проецируется **A**.

3. При $i \rightarrow L$, модель (5) позволяет учитывать особенности речевосприятия через (3), а речеобразования через (4).

Теперь, исходя из (5), модель сингулярного оценивания ЧОТ можно представить в следующем концептуальном виде (рисунок 2):

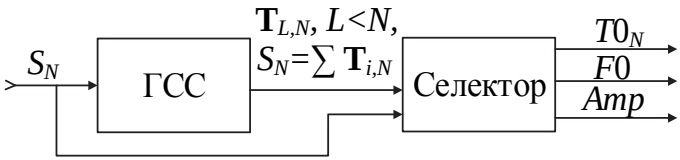


Рис. 2. Концептуальная модель сингулярного оценивания ЧОТ: S_N – входной сигнал; $\mathbf{T}_{L,N}$ – спектр временных рядов; ГСС – генератор сингулярного спектра; S_N – входной сигнал; $T0$ – трек основного тона; $F0$ – ЧОТ; *Amp* – амплитуда

1) средство генерации сингулярного спектра речевого сигнала, в котором входные данные – это фонемный ряд S_N, а выходные данные – это спектр временных рядов $\mathbf{T}_{L,N}$;

2) средство выбора спектральной составляющей соответствующей частоте основного тона речи, в котором входные данные – это спектр временных рядов $\mathbf{T}_{L,N}$, а выходные данные – это частота основного тона речи $F0$, средняя амплитуда Amp и квазигармоническая составляющая основного тона речи $T0$.

Численная реализация модели сингулярного оценивания ЧОТ речи выражается в системе:

$$
\begin{cases}
\mathbf{A} = [S_{i\text{-}1},...,S_{i+L\text{-}1}]^{\mathrm{T}}, 1 \le i \le K, K = N - L + 1; \\
\mathbf{C} = \mathbf{A}\mathbf{A}^{\mathrm{T}}; \\
(\mathbf{U}_{\mathrm{C}}, \mathbf{D}_{\mathrm{C}}) = \mathrm{Eigens}(\mathbf{C}); \\
\mathbf{V}_{\mathrm{A}}{}^{\mathrm{T}} = \mathrm{D}_{\mathrm{C}}{}^{-1}\mathbf{U}_{\mathrm{C}}{}^{\mathrm{T}}\mathbf{A}; \\
(3); \\
f_n = \\
\dfrac{p}{N\Delta t}, p = \{k, \left\| [\dfrac{1}{N}\sum_{j=1}^{N} \mathbf{T}_j^{<n>} e^{\frac{-2\pi \mathbf{i}}{N}kj}]_k \right\| \subseteq MAX, k = \overline{1,N}\}, \\
n = \overline{1,L}; \\
f_j = f_n \in [f_{\min} \le f_n \le f_{\max}], n = \overline{1,L}, \\
j = 0,1,...,K < L; \\
f_0 = f_{j=\text{нкчот}} = f_j \in \{\min(f_j), 2\min(f_j), \ldots, M\min(f_j)\}\}, \\
j = \overline{1,K}; \\
T0_n = T_{j=\text{нкчот},n}, n = \overline{1,N}; \\
F0 = \dfrac{1}{m-1}\sum_{i=1}^{m} \dfrac{1}{(k_{i-1} - k_i)\Delta t}, \\
k_i = \{n, T0_n \subset MAX, n = \overline{0,N-1}\}, i = \overline{1,m}; \\
Amp = \dfrac{1}{m}\sum \max(T0_n), n = 1,2,...,m.
\end{cases}
$$

где: S_N – исходный временной ряд;
N – длина ряда;
L – размер спектрального окна;
$\mathbf{A}$ – траекторная (Ганкелева / H. Hankel matrix) матрица наблюдений [11];
$\mathbf{C}$ – бисимметричная матрица;
$\mathbf{U}_{\mathrm{C}}$ – левая сингулярная матрица поворота;

$\mathbf{V}_A{}^T$ – правая сингулярная матрица поворота;
$\mathbf{u}^{<n>}$ – левый сингулярный вектор;
$\mathbf{v}^{<n>}$ – правый сингулярный вектор ($\mathbf{v}^{<n>}=\mathbf{x}^{<n>}\in\mathbf{V}$);
$\mathbf{D}$ – диагональная матрица, состоящая из собственных значений λ_i бисимметричной матрицы $\mathbf{C}$, при условии:

$$\mathbf{D}_C = \mathrm{diag}\{\lambda_0,\lambda_1,\lambda_2,...,\lambda_{L-1}\},\ \lambda_0 < \lambda_1 <,...,< \lambda_{L-1};$$

$\mathbf{T}_i^n$ – спектр временных рядов (квазигармонический спектр);
Eigens – функция поиска собственных чисел;
НКЧОТ – номер компоненты с частотой основного тона;
$\mathbf{T}_{j=\text{нкчот},N}$ – активация квазигармоники с НКЧОТ;
f_n – одномерное, частотное представление временного спектра $\mathbf{T}_i^n$ при условии, что $f_0\in[f_{\min}, f_{\max}]$, где f_0 – искомая частота основного тона такая, что:

$$f_0 \in \{\min(f_i),2\min(f_i),\ldots,M\min(f_i)\}$$

наименьшая кратная величина частоты;
p – индекс элемента в ряде T_i^n , соответствующий максимальной амплитуде от преобразований Фурье в n-й квазигармонике;
Δt – величина обратная частоте дискретизации;
$T0_N$ – временной ряд, соответствующий квазигармонике с частотой основного тона речи;
$F0$ – средняя частота основного тона речи такая, что:

$$F0 = \frac{f_0^1 + f_0^2 + ... + f_0^m}{m-1},$$

где $(m\text{-}1)$ – число обратных величин равных периодам умещающихся в ряде $T0_N$ (f_0^i – локальная частота тона):

$$f_0^1 + f_0^2 + ... + f_0^m = \frac{1}{(k_2-k_1)\Delta t} + \frac{1}{(k_3-k_2)\Delta t} + \ldots$$

$$+\frac{1}{(k_m-k_{m-1})\Delta t} = \sum_{i=1}^{m}\frac{1}{(k_i-k_{i-1})\Delta t},$$

где k_i – номер индекса в точке максимума:

$$k_i = \{n, T0_n \subset max,\ n = \overline{0,N-1}\},\ i = \overline{1,m};$$

Amp – средняя амплитуда квазигармоники основного тона речи.

3. Описание программной реализации сингулярного оценивания частоты основного тона. На примере модели "черный ящик" [18–20] рассмотрим программный комплекс сингулярного оценивания ЧОТ, состоящий из 10-ти программно реализованных модулей (рисунок 3):

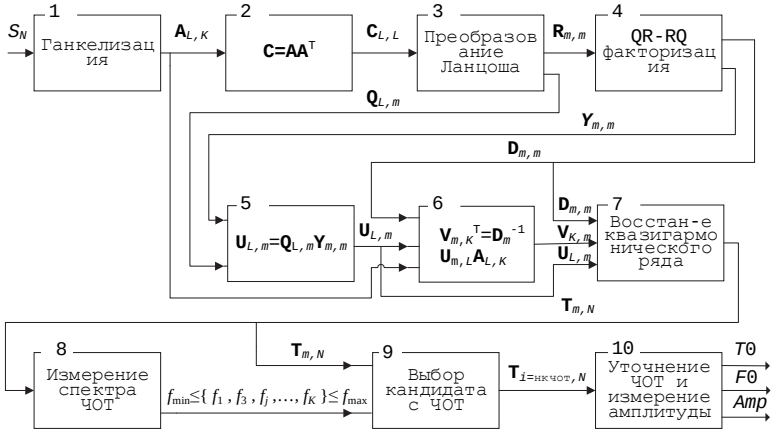


Рис. 3. Структура программного комплекса на уровне блоков

1) модуль Ганкелизации речевого сигнала S_N для получения траекторной матрицы $\mathbf{A}_{L,K}$;

2) модуль вычисления ковариационной матрицы $\mathbf{C}_{L,L}=\mathbf{A}\mathbf{A}^{\mathrm{T}}$;

3) модуль преобразований Ланцоша для вычисления трехдиагональной матрицы Релея $\mathbf{R}$ размерностью $m\times m$ и ортонормированного базиса подпространства Крылова $\mathbf{Q}_{L,m}$ [21, 22];

4) модуль QR факторизации для отыскания собственных пар ($\mathbf{y}^{<n>}\in\mathbf{Y}_{m,m}$, $\lambda_n\in\mathbf{D}$) матрицы Релея $\mathbf{R}_{m,m}$, где: $\mathbf{R}=\mathbf{Y}\,\mathbf{D}\,\mathbf{Y}^{\mathrm{T}}$, $\mathbf{Y}$–матрица собственных векторов матрицы $\mathbf{R}$, $\mathbf{D}$ – матрица собственных значений матрицы $\mathbf{R}$;

5) модуль вычисления первых m собственных векторов $\mathbf{u}^{<n>}\in\mathbf{U}$ (поиск матричной пары Ритца ($\mathbf{U}_{L,m}$, $\mathbf{D}_{m,m}$), где $\mathbf{U}_{L,m}=\mathbf{Q}_{L,m}\mathbf{Y}_{m,m}$ – матрица, состоящая из векторов Ритца);

6) модуль вычисления первых m собственных векторов $\mathbf{v}^{<n>}\in\mathbf{V}$ (матрицы главных компонент) траекторной матрицы $\mathbf{A}$, порождаемых ее строками:

$$\mathbf{V}_{m,K}^{\mathrm{T}} = \mathbf{D}_{m}^{-1}\mathbf{U}_{m,L}\mathbf{A}_{L,K};$$

7) модуль реконструкции первых m компонент квазигармонического спектра $\mathbf{T}_{m,N}$ речевого сигнала;

8) модуль измерения частоты квазигармонического спектра $\mathbf{T}_{m,N}$ (блок измерения частоты временного спектра);

9) модуль выбора кандидата с ЧОТ (блок выбора номера компоненты с частотой основного тона);

10) модуль уточнения ЧОТ и измерения амплитуды (блок вычисления частоты и амплитуды основного тона).

В соответствии с концептуальной моделью сингулярного измерителя, блоки 1-7 описывают генератор сингулярного спектра (ГСС), а блоки 8-10 описывают средство выбора квазигармонической составляющей основного тона и дальнейшего уточнения частоты и амплитуды основного тона (Селектор).

Таким образом, программная реализация сингулярного измерителя ЧОТ включает:

1. Конструктор класса генератора сингулярного спектра, реализованный в качестве функции (рисунок 4):

$$T=ssg(S, N, L, m),$$

где: S – массив данных, содержащий исходный фонемный ряд S_N;
N – размер массива S;
L – размер окна анализа;
m – число квазигармонических составляющих;
T – массив данных, содержащий квазигармонический спектр.

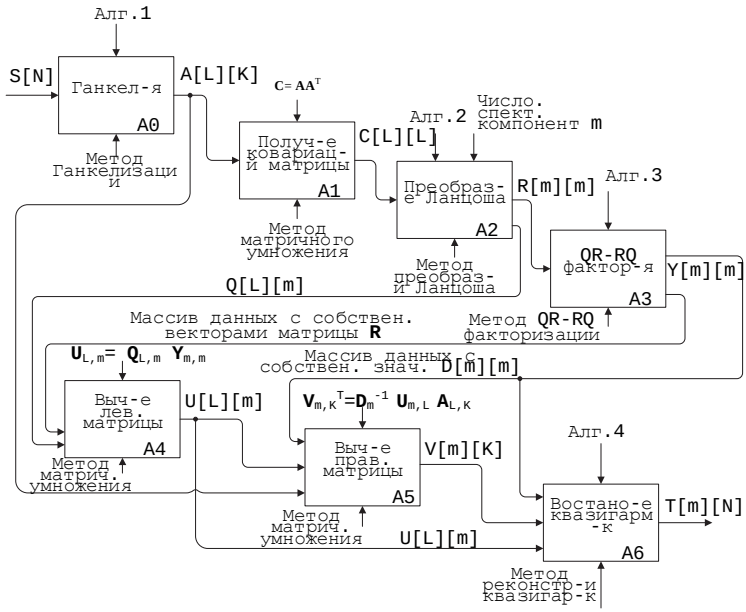


Рис. 4. Программная реализация генератор сингулярного спектра на уровне модели IDEF0

2. Модуль преобразований Ланцоша, реализованный в качестве метода класса ssg:

$$(R, Q)= \text{Lanczos}(C, RS, CS),$$

где: C – массив данных, содержащий ковариационную матрицу;
RS, CS – параметры, задающие размеры ковариационной матрицы **C**;
R – массив данных, содержащий трехдиагональную симметричную матрицу $\mathbf{R}_{m,m}$;
Q – массив данных, содержащий векторы Ланцоша матрицы $\mathbf{Q}_{L,m}$.

3. Модуль QR факторизации, реализованный в качестве метода класса ssg:

$$(D, Y)=qr(a, b, RS),$$

где:
a – массив данных, содержащий элементы трехдиагональной матрицы **R**, расположенных на главной диагонали;
b – массив данных, содержащий элементы трехдиагональной матрицы **R**, расположенных над главной диагональю;
RS –входной параметр, задающий размер массива a (количество спектральных компонент RS=*m*);
Y – массив данных, содержащий собственные векторы матрицы **R**.

4. Конструктор класса селектора, реализованный в качестве функции (рисунок 5):

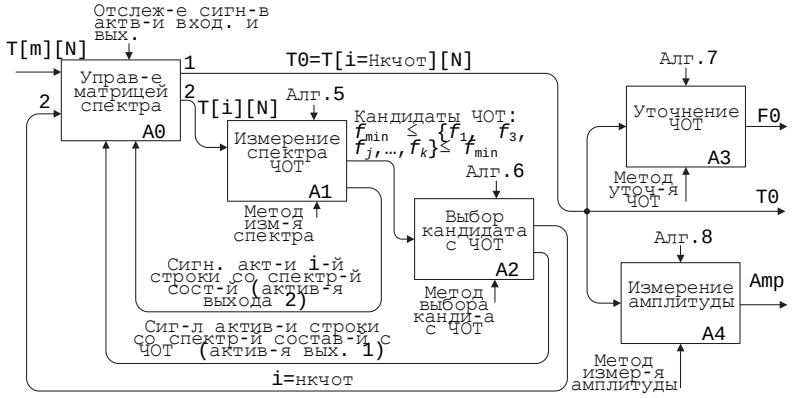


Рис. 5. Программная реализация селектора на уровне модели IDEF0

Selector(S, T, T0, F0, Amp, m, N),

где: T – массив данных, содержащий квазигармонический спектр;
T0 – массив данных, содержащий квазигармонику основного тона;
F0 – переменная, содержащая ЧОТ;
Amp – переменная, содержащая среднюю амплитуду квазигармоники основного тона.

4. Тестирование программной реализации сингулярного оценивания частоты основного тона. Рассмотрим общий вид работы программной реализации сингулярного оценивания ЧОТ. На вход программы подаются данные в виде фонемного ряда S, который выступает в качестве входного параметра для инициализации класса ГСС. Конструктор класса ГСС вызывает методы, в которых реализованы алгоритмы сингулярного спектрального анализа. В процессе работы вызываемых методов, осуществляется преобразование фонемного ряда, содержащегося в массиве данных S, в спектр квазигармоник, содержащихся в двухмерном массиве данных T. Массив данных T выступает в качестве входного аргумента при инициализации класса селектора. Конструктор класса селектора вызывает методы выбора квазигармонической составляющей с ЧОТ, и методы расчета его параметров. На выходе программы данные, соответствующие параметрам основного тона T0, Amp, F0.

Оценка временных характеристик сингулярного оценивания ЧОТ проводилась на персональном компьютере на базе процессора Intel i5 3.1GHz и мобильном устройстве связи на базе процессора Apple A6 1.7GHz (таблица 1). В качестве положительного критерия временных характеристик оценивания ЧОТ принималась работа программы в режиме реального времени. Под режимом реального времени понимается время сингулярного оценивания ЧОТ меньшее, чем сам кадр анализа. В качестве входных данных выбирались фонемные ряды гласных звуков русской речи, мужского и женского диктора, длительностью 32мс. В таблице 1 параметр G задает количество спектральных составляющих, которые необходимо найти, а ε задает достаточную ошибку округления для сингулярных чисел. Для уменьшения латентности анализа подбираются соответствующие параметры G и ε. Результаты тестирования временных характеристик показывают, что время оценивания ЧОТ (выполнения программы) как для ПК (Intel i5

3.1Ghz), так и для мобильного устройства связи (Apple A6 1.7Ghz) не превышает заданного начальным условием.

Таблица 1. Временные характеристики оценивания ЧОТ

Диктор № 2, пол: Ж, *G*=32									
Фонема	нк чот	Intel i5 3.1GHz				Apple A6 1.7GHz			
		Время (мс)	ЧОТ (Гц) при ε= 0,00001	Время (мс)	ЧОТ (Гц) при ε= 0,0001	Время (мс)	ЧОТ (Гц) при ε= 0,00001	Время (мс)	ЧОТ (Гц) при ε= 0,0001
[ɑ]	1	15	199,804	9	199,804	28	199,804	18	199,800
[e]	2	17	195,047	7	195,047	27	195,047	20	195,000
[ɵ]	2	15	199,804	9	199,804	29	199,804	25	199,800
[i]	1	16	204,800	8	204,800	27	204,800	22	204,800
[o]	2	17	210,051	10	210,051	26	210,051	21	204,800
[u]	1	18	215,578	6	215,578	29	215,578	20	215,570
[ɨ]	1	15	204,800	7	204,800	28	204,800	20	204,800
[ɪ]	2	18	199,804	9	199,804	30	199,804	18	199,800
[ʉ]	1	17	204,800	10	204,800	38	204,800	21	204,800
[æ]	1	16	186,181	7	186,181	27	186,181	19	186,180

Оценка точности сингулярного оценивания ЧОТ проводилась при следующих условиях:

1. Для известных алгоритмов RAPT, YIN, SWIPE', SHS, AC-P, AC-S, ANAL, CC, CEP, ESRPD, SHR, TEMPO [4-7, 23-29] и сингулярного оценивания ЧОТ (SEPT – Singular Estimation Pitch Tracking) рассматривался процент грубых ошибок GPE (gross pitch errors) [9]. Величина GPE показывает отношение количества анализируемых фреймов с отклонением полученной оценки ЧОТ более чем на ±20% от реального значения ЧОТ к общему числу вокализированных фреймов:

$$GPE(\%) = \frac{N_{\mathrm{GPE}}}{N_{\mathrm{V}}} 100,$$

где: N_{GPE} – число фреймов с отклонением полученной оценки более чем на ±20% от настоящего значения ЧОТ;
N_{V} – общее число вокализированных фреймов.

На первый взгляд 20%-я погрешность ошибки кажется слишком большой, но, учитывая, что большинство ошибок, допускаемых алгоритмами при оценивании ЧОТ варьируется в пределах октавы, то выбор такой погрешности можно считать обоснованным.

2. Доступ к известным алгоритмам осуществлялся с помощью программного обеспечения SFS [30], Praat [31], Straight [32], Aubio [33], Festival [34], SPE [35], (таблица 2).

3. В качестве анализируемого материала были выбраны речевые базы Disordered Voice Database (DVD) [36], Keele Pitch Database (KPD) [37] и Paul Bagshaw's Database (PBD) [38].

Если принять, что реализация всех алгоритмов выполнена в соответствии с их оригинальным описанием [4-7, 23-29], то при использовании идентичных входных данных (речевых фрагментов из выбранных баз) и единого аппаратного обеспечения (ПК на базе Intel i5 3.1GHz) можно считать, что сравнение алгоритмов проводились в идентичных условиях.

Таблица 2. Перечень алгоритмов оценивания ЧОТ и доступ к ним

Метод	Алгоритм	Программа (библиотека)	Доступ (функция)
Автокорреляционный	AC-S	SFS	fxac
Автокорреляционный	ANAL	SFS	fxanal
Кепстральный	CEP	SFS	fxcep
Кросскорреляционный	RAPT	SFS	fxrapt
Автокорреляционный	AC-P	PRAAT	ac
Кросскорреляционный	CC	PRAAT	cc
Спектральный	SHS	PRAAT	shs
Кросскорреляционный	ESRPD	FESTIVAL	pda
Спектральный	SHR	MATLABCENT.	shrp
Корреляционный	SWIPE'	SPE	Swipe
Спектральный	TEMPO	STRAIGHT	exstraightsource
Автокорреляционный	YIN	AUBIO	Aubiopitch

Таблица 3 показывает характеристику GPE для несортированных образцов вокализированных сегментов речи, выбранных из базы DVD (рисунок 6).

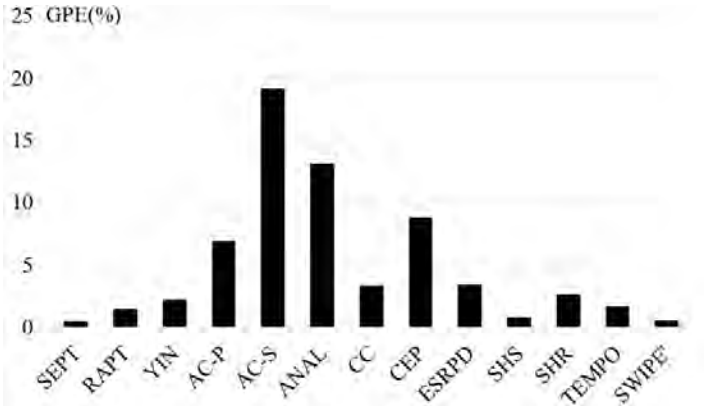


Рис. 6. Средний процент грубых ошибок оценки ЧОТ, допускаемый известными алгоритмами и SEPT

Таблица 3. Оценка грубых ошибок для натуральной речи

	Процент грубых ошибок - GPE (%)			
Алгоритм	База PBD	База KPD	База DVD	Среднее
SEPT	0,11	0,65	0,74	0,49
RAPT	0,78	1,08	2,70	1,52
YIN	0,35	1,43	4,90	2,22
AC-P	0,72	3,01	17,02	6,91
AC-S	8,90	7,40	41,18	19,16
ANAL	0,81	2,70	36,05	13,18
CC	0,45	3,65	6,03	3,37
CEP	6,20	4,23	16,07	8,83
ESRPD	1,35	3,99	5,00	3,44
SHS	0,16	1,03	1,34	0,84
SHR	0,71	1,56	5,80	2,69
TEMPO	0,33	1,97	2,91	1,73
SWIPE’	0,14	0,87	0,80	0,60

Таблица 4 показывает характеристику GPE в зависимости от пола диктора для баз PBD и KPD, т.к. для них имеются контрольные оценочные значения ЧОТ, полученные с помощью Ларинографа.

Таблица 4. Оценка грубых ошибок по полу

	Процент грубых ошибок - GPE (%)		
Алгоритм	Мужчины	Женщины	Среднее
SEPT	0,32	2,10	1,21
RAPT	0,45	2,81	1,63
YIN	1,19	3,12	2,16
AC-P	2,25	3,55	2,90
AC-S	3,17	9,40	6,29
ANAL	1,41	5,73	3,57
CC	2,54	4,43	3,49
CEP	2,00	4,17	3,09
ESRPD	3,20	3,79	3,50
SHS	0,62	2,39	1,51
SHR	0,68	3,57	2,13
TEMPO	0,75	2,98	1,87
SWIPE’	0,40	2,32	1,36
Среднее	1,97	3,81	2,89

На первый взгляд, величина GPE показывает степень робастности оценивания ЧОТ, так как, по сути, показывает процент допущенных ошибок каждым алгоритмов в процессе оценивания, но с другой стороны по данной величине можно судить о степени точности оценки

ЧОТ. Так, например, можно сказать, что для несортированной базы DVD, SEPT на 20% робастен по отношению к SWIPE', так как более точнее осуществляет оценку ЧОТ за счет сингулярной модели речевого сигнала, которая позволяет рассматривать речеобразующий тракт как систему акустических резонаторов, в которой параметрами выступают собственные значения и собственные векторы, содержащие информацию о структуре речевого сигнала с учетом нестационарных амплитуд, периодов и фаз гармоник, входящих в его состав.

5. Заключение. Время оценивания ЧОТ для 100 несортированных вокализованных образцов речи для ПК, при заданном ε=0,00001, не превышало 20мс. Таким образом, можно заключить, что сингулярное оценивание ЧОТ можно использовать в приложениях реального времени, где задержка в 20мс может быть допустимой. Результаты эксперимента показывают, что сингулярный метод оценивания ЧОТ более робастен по отношению к известным аналогам, а значит более точнее оценивает ЧОТ.

Литература

1. *Голубинский А.Н.* Оценка частоты основного тона речевого сигнала при априори неизвестных амплитудах и начальных фазах полигармонического несущего колебания // Вестник Воронежского института МВД России. 2010. № 3. С. 110–117.
2. *Ронжин А.Л., Басов О.О.* Определение степени алкогольной интоксикации человека на основе автоматического анализа речи // Вестник Московского университета МВД России. 2015. № 5. С. 216–220.
3. *Meshcheryakov R.V., Balatskaya L.N., Choinzonov E.L., Chizevskaya S.Yu., Kostyuchenko E.U.* Software for Assessing Voice Quality in Rehabilitation of Patients after Surgical Treatment of Cancer of Oral Cavity, Oropharynx and Upper Jaw // Proceedings of 15th International Conference SPECOM 2013. Pilsen. Czech Republic. 2013. pp 294–301.
4. *Talkin D.* A Robust Algorithm for Pitch Tracking (RAPT) // Speech Coding & Synthesis. 1995. pp-495–518.
5. *Cheveigne A., Kawahara H.* YIN, a fundamental frequency estimator for speech and music // Jour. Acoust. Soc. Am. 2002. vol. 111. no. 4. pp. 1917–1930.
6. *Camacho A., Harris J.G.* A sawtooth waveform inspired pitch estimator for speech and music // Journal Acoust. Soc. Am. 2008. vol. 123. no. 4. pp. 1638–1652.
7. *Hermes D.J.* Measurement of pitch by subharmonic summation // Jour. Acoust. Soc. Am. 1988. vol. 83. pp. 257–264.
8. *Rabiner L.R., Schafer R.W.* Digital processing of speech signals // Prentice Hall. 1978.
9. *Azarov E., Vashkevich M., Petrovsky A.* Instantaneous pitch estimation based on RAPT framework // Proceedings of the 20th European Signal Processing Conference (EUSIPCO). Bucharest. 2012. pp. 2787–2791.
10. *Basov O.O., Ronzhin A.L., Budkov V.Yu.* Optimization of Pitch Tracking and Quantization // Proc. SPECOM-2015. LNAI 9319. 2015. pp. 65–72.
11. *Basov O.O., Ronzhin A.L., Budkov V.Yu.,* Saitov *I.A.* Method of Defining Multimodal Information Falsity for Smart Telecommunication Systems // Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer. St. Petersburg. Russia. 2015. LNCS 9247. pp. 163–173.

12. *Golyandina N., Zhigljavsky A.* Singular Spectrum Analysis for time series // Springer Science & Business Media. 2013.
13. *Tony F.C.* An Improved Algorithm for Computing the Singular Value Decomposition // ACM Transaction on Mathematical Software. 1982. vol. 8. no. 1. pp. 72–83.
14. *Азаров И.С., Вашкевич М.И., Лихачев Д.С., Петровский А.А.* Изменение частоты основного тона речевого сигнала на основе гармонической модели с нестационарными параметрами // Труды СПИИРАН. 2014. Вып. 32. С.5–26.
15. *Бондаренко В.П., Коцубинский В.П., Мещеряков Р.В.* Нестационарные модели в обработке речевых сигналов // Акустика речи. Медицинская и биологическая акустика. Архитектурная и строительная акустика и вибрации. Сб. трудов XVIII сессии Российского акустического общества. М.: ГЕОС. 2006. Т.3. С. 8–11.
16. *Вольф Д.А.* Спектральная теорема для решения частичной проблемы собственных чисел степенным методом в задачах сингулярного спектрального анализа речи // Системы управления и информационные технологии. 2014. №3.1(57). С. 129–135.
17. *Агаев Р.П., Чеботарев П.Ю.* Метод проекции в задаче о консенсусе и регуляризованный предел степеней стохастической матрицы //Автоматика и телемеханика. 2011. №. 12. С. 38–59.
18. *Налимов В.В.* Теория эксперимента // М.: Наука. 1971. 208 с.
19. *Силич В.А., Комагоров В.П., Савельев А.О.* Принципы разработки системы мониторинга и адаптивного управления разработкой «интеллектуального» месторождения на основе постоянно действующей геологотехнологической модели // Известия Томского политехнического университета. 2013. Т. 323. №. 5. С. 94–100.
20. *Силич В. А., Силич М.П., Аксенов С.В.* Алгоритм построения нечеткой системы логического вывода Мамдани, основанный на анализе плотности обучающих примеров //Доклады томского государственного университета систем управления и радиоэлектроники. 2013. №. 3(29). С. 76–82.
21. *Parlett B. N.* The symmetric eigenvalue problem // Englewood Cliffs. NJ: Prentice-Hall. 1980. vol. 7.
22. *Knizhnerman L., Simoncini V.* A new investigation of the extended Krylov subspace method for matrix function evaluations // Numerical Linear Algebra with Applic. 2010. vol. 17. no. 4. pp. 615–638.
23. *Boersma P.* Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound // Proceedings of the institute of phonetic sciences. 1993. vol. 17. no. 1193. pp. 97–110.
24. *Secrest B. G., Doddington G. R.* An integrated pitch tracking algorithm for speech systems // Acoustics, Speech, and Signal Processing. IEEE International Conference on ICASSP'83. 1983. vol. 8. pp. 1352–1355.
25. *Noll A. M.* Cepstrum pitch determination // The journal of the acoustical society of America. 1967. vol. 41. no. 2. pp. 293–309.
26. *Bagshaw P. C., Hiller S. M., Jack M. A.* Enhanced pitch tracking and the processing of F0 contours for computer and intonation teaching // Proc. Europe-an Conf. on Speech Comm. (Eurospeech). 1993. pp. 1003–1006.
27. *Medan Y., Yair E., Chazan D.* Super resolution pitch determination of speech signals // IEEE Trans. Signal Process. 1991. vol. 39. pp. 40-48.
28. *Sun X.* A pitch determination algorithm based on subharmonic-to-harmonic ratio // The 6th International Conference of Spoken Language Processing. 2000. pp. 676–679.
29. *Kawahara H., Katayose H., de Cheveigne A., Patterson R. D.* Fixed Point Analysis of Frequency to Instantaneous Frequency Mapping for Accurate Esti-mation of F0 and Periodicity // Proc. EUROSPEECH. 1999. vol. 99. Issue 6. pp. 2781–2784.

30. Speech Filing System (SFS) // UCL Psychology & Language sciences Faculty of Brain Sciences. 2015. URL: http://www.phon.ucl.ac.uk/resource/sfs/ (дата обращения: 17.09.2015).
31. Praat // Phonetic Sciences. Amsterdam. 2015. URL: http://www.fon.hum.uva.nl/praat/download_win.html (дата обращения: 17.09.2015).
32. Straight // GitHub. 2015. URL: https://github.com/shuaijiang/STRAIGHT (дата обращения: 17.09.2015).
33. Aubio //Aubio. 2015. URL: http://aubio.org/download (дата обращения: 17.09.2015).
34. Festival // The Festival Speech Synthesis System. 2015. URL: http://www.cstr.ed.ac.uk/projects/festival/download.html (дата обращения: 17.09.2015).
35. SWIPE' pitch estimator. 2015. URL: https://github.com/kylebgorman/swipe (дата обращения: 17.09.2015).
36. Disordered Voice Database. 2015. URL: http:// http://kayelemetrics.com (дата обращения: 17.09.2015).
37. Keele Pitch Database. 2015. URL: http://www.icocla.it/keele.html (дата обращения: 20.03.2015).
38. Paul Bagshaw's Database. 2015. URL: http://www.cstr.ed.ac.uk/research/projects/fda (дата обращения: 17.09.2015).

References

1. Golubinsky A.N. [Pitch frequency estimation of a speech signal at a priori unknown amplitudes and initial phases of polyharmonical carrying oscillation]. *Vestnik Voronezhskogo instituta MVD Rossii – Vestnik of Voronezh Institute of the Ministry of the Interior of Russia*. 2010. vol.3. pp. 110–117. (In Russ.).
2. Ronzhin A.L., Basov O.O. [Detection of alcohol intoxication degree based on automatic speech analysis]. *Vestnik Moskovskogo universiteta MVD Rossii – Herald of Moscow University of the MIA of Russia*. 2015. no. 5. pp. 216–220. (In Russ.).
3. Meshcheryakov R.V., Balatskaya L.N., Choinzonov E.L., Chizevskaya S.Yu., Kostyuchenko E.U. Software for Assessing Voice Quality in Rehabilitation of Patients after Surgical Treatment of Cancer of Oral Cavity, Oropharynx and Upper Jaw. Proceedings of 15th International Conference SPECOM 2013. Pilsen. Czech Republic. 2013. pp 294–301.
4. Talkin D. A Robust Algorithm for Pitch Tracking (RAPT). *Speech Coding & Synthesis*. 1995. pp-495–518.
5. Cheveigne A., Kawahara H. YIN, a fundamental frequency estimator for speech and music. *Jour. Acoust. Soc. Am.* 2002. vol. 111. no. 4. pp. 1917–1930.
6. Camacho A., Harris J.G. A sawtooth waveform inspired pitch estimator for speech and music. *Journal Acoust. Soc. Am.* 2008. vol. 123. no. 4. pp. 1638–1652.
7. Hermes D.J. Measurement of pitch by subharmonic summation. *Jour. Acoust. Soc. Am.* 1988. vol. 83. pp. 257–264.
8. Rabiner L.R., Schafer R.W. Digital processing of speech signals. Prentice Hall. 1978.
9. Azarov E., Vashkevich M., Petrovsky A. Instantaneous pitch estimation based on RAPT framework. Proceedings of the 20th European Signal Processing Conference (EUSIPCO). Bucharest. 2012. pp. 2787–2791.
10. Basov O.O., Ronzhin A.L., Budkov V.Yu. Optimization of Pitch Tracking and Quantization. Proc. SPECOM-2015. LNAI 9319. 2015. pp. 65–72.
11. Basov O.O., Ronzhin A.L., Budkov V.Yu., Saitov I.A. Method of Defining Multimodal Information Falsity for Smart Telecommunication Systems. Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer. St. Petersburg. Russia. 2015. LNCS 9247. pp. 163–173.

12. Golyandina N., Zhigljavsky A. Singular Spectrum Analysis for time series. Springer Science & Business Media. 2013.
13. Tony F.C. An Improved Algorithm for Computing the Singular Value Decomposition. ACM Transaction on Mathematical Software. 1982. vol. 8. no. 1. pp. 72–83.
14. Azarov E., Vashkevich M., Likhachov D., Petrovsky A. Pitch modification of speech signal using harmonic model with time-varying parameters. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2014. vol. 32. pp.5–26. (In Russ.).
15. Bondarenko V.P., Kotsubinsky V.P., Meshcheryakov R.V. [Non-stationary models in speech signal processing]. *Acistica rechi. Medicinskaja i biologicheskaja akustika. Arhitekturnaja i stroitel'naja akustika i vibracii. Sbornik trudov XVIII sessii Russain acoustic society* [Acoustics of speech. Medical and biological acoustics. Architectural and building acoustics and vibration. Proceedings of the 17th Session of the Russian Acoustical Society. M: GEOS. 2006. vol. 3. pp. 8–11. (In Russ.).
16. Volf D.A. [The use of spectral theorem for power-iteration solution of a partial eigenvalue problem in singular spetrum analysis of speech]. *Sistemy Upravleniia i Informatsionnye Tekhnologii – Control Systems and Information Technology*. 2014. no. 3.1(57). pp. 129–135. (In Russ.).
17. Agaev R.P., Chebotarev P.Yu. [The projection method for reaching consensus and the regularized power limit of a stochastic matrix]. *Avtomatika i Telemekhanika – Automation and Remote control*. 2011. vol. 12. pp. 38–59. (In Russ.).
18. Nalimov V.V. *Teorija jeksperimenta* [The theory of experiment]. 1971. 208 p. (In Russ.).
19. Silich V. A., Komagorov V. P., Savelev A. O. [Principles of developing the system of monitoring and adaptive controlling the intelligent oil field study based on permanent geological and technological models]. *Izvestija Tomskogo politehnicheskogo universiteta – Tomsk Polytechnic University*. 2013. vol. 323. no. 5. pp. 94–100. (In Russ.).
20. Silich V. A., Yampolsky V.Z., Savelyev A.O., Komagorov V.P., Alekseev A.A., Grebenshchikov S.A. [A Mamdani-type fuzzy system construction algorithm based on training vectors density analysis]. *Doklady tomskogo gosudarstvennogo universiteta sistem upravlenija i radiojelektroniki – Proceedings of Tomsk State University of Control Systems and Radioelectronics*. 2013. vol. 3(29). pp. 76-82. (In Russ.).
21. Parlett B. N. The symmetric eigenvalue problem. Englewood Cliffs. NJ: Prentice-Hall. 1980. vol. 7.
22. Knizhnerman L., Simoncini V. A new investigation of the extended Krylov subspace method for matrix function evaluations. *Numerical Linear Algebra with Applic*. 2010. vol. 17. no. 4. pp. 615–638.
23. Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. Proceedings of the institute of phonetic sciences. 1993. vol. 17. no. 1193. pp. 97–110.
24. Secrest B.G., Doddington G.R. An integrated pitch tracking algorithm for speech systems. Acoustics, Speech, and Signal Processing. IEEE International Conference on ICASSP'83. 1983. vol. 8. pp. 1352–1355.
25. Noll A.M. Cepstrum pitch determination. *The journal of the acoustical society of America*. 1967. vol. 41. no. 2. pp. 293–309.
26. Bagshaw P.C., Hiller S.M., Jack M.A. Enhanced pitch tracking and the processing of F0 contours for computer and intonation teaching. Proc. Europe-an Conf. on Speech Comm. (Eurospeech). 1993. pp. 1003–1006.
27. Medan Y., Yair E., Chazan D. Super resolution pitch determination of speech signals. *IEEE Trans. Signal Process.* 1991. vol. 39. pp. 40-48.
28. Sun X. A pitch determination algorithm based on subharmonic-to-harmonic ratio. The 6th International Conference of Spoken Language Processing. 2000. pp. 676–679.

29. Kawahara H., Katayose H., de Cheveigne A., Patterson R.D. Fixed Point Analysis of Frequency to Instantaneous Frequency Mapping for Accurate Esti-mation of F0 and Periodicity. Proc. EUROSPEECH. 1999. vol. 99. Issue 6. pp. 2781–2784.
30. Speech Filing System (SFS). UCL Psychology & Language sciences Faculty of Brain Sciences. Available at: http://www.phon.ucl.ac.uk/resource/sfs/ (accessed 17.09.2015).
31. Praat. Phonetic Sciences, Amsterdam. 2015. Available at: http://www.fon.hum.uva.nl/praat/download_win.html (дата обращения: 17.09.2015).
32. Straight. GitHub. Available at: https://github.com/shuaijiang/STRAIGHT (accessed 17.09.2015).
33. Aubio. Aubio. Available at: http://aubio.org/download (accessed 17.09.2015).
34. The Festival Speech Synthesis System. Available at: http://www.cstr.ed.ac.uk/projects/festival/download.html (accessed 17.09.2015).
35. SPE. SWIPE' pitch estimator. Available at: https://github.com/kylebgorman/swipe (accessed 17.09.2015).
36. Disordered Voice Database. Available at: http:// http://kayelemetrics.com (accessed 17.09.2015).
37. Keele Pitch Database. Available at: http://www.icocla.it/keele.html (accessed 20.03.2015).
38. Paul Bagshaw's Database. Available at: http://www.cstr.ed.ac.uk/research/projects/fda (accessed 17.09.2015).

Вольф Данияр Александрович — аспирант кафедры комплексной информационной безопасности электронно-вычислительных систем, Томский государственный университет систем управления и радиоэлектроники (ТУСУР). Область научных интересов: системный анализ, сингулярный анализ, программирование, моделирование. Число научных публикаций — 16. runsolar@mail.ru; пр. Ленина, 40, Томск, 634050; р.т.: +7(3822)900-111, Факс: +7(3822)900-111.

Volf Daniiar Aleksandrovich — Ph.D. student of complex security of electronic-computing systems department, Tomsk State University of Control Systems and Radioelectronics (TUSUR). Research interests: speech analysis, speech recognition, signal analisys. The number of publications — 16. runsolar@mail.ru; 40, Lenin-avenue Tomsk, 634050, Russia; office phone: +7(3822)900-111, Fax: +7(3822)900-111.

Мещеряков Роман Валерьевич — д-р техн. наук, доцент, профессор кафедры комплексной информационной безопасности электронно-вычислительных систем, Томский государственный университет систем управления и радиоэлектроники (ТУСУР). Область научных интересов: системный анализ, информационная безопасность, вопросы обработки информации в интеллектуальных системах, особое внимание уделяется вопросам создания информационно-безопасных систем. Число научных публикаций — 247. mrv@ieee.org; пр. Ленина, 40, Томск, 634050; р.т.: +7(3822)900111, Факс: +7 (3822) 900-111.

Meshcheryakov Roman Valerievich — Ph.D., Dr. Sci., professor, professor of complex security of electronic-computing systems department, Tomsk State University of Control Systems and Radioelectronics (TUSUR). Research interests: speech analysis, speech recognition, medical technology, information security. The number of publications — 247. mrv@ieee.org; 40, Lenin-avenue Tomsk, 634050, Russia; office phone: +7(3822)900111, Fax: +7(3822)900111.

Поддержка исследований. Работа выполнена в рамках государственного задания Томского государственного университета систем управления и радиоэлектроники (проект № 3657).

Acknowledgements. Tomsk State Unuversity of Control Systems and Radioelectronics (project 3657).

РЕФЕРАТ

Вольф Д.А. Мещеряков Р.В. **Модель и программная реализация сингулярного оценивания частоты основного тона речевого сигнала.**

В статье рассматривается сингулярная модель речевого сигнала, которая позволяет рассматривать речеобразующий тракт как систему акустических резонаторов, в которой параметрами выступают собственные значения и собственные векторы, содержащие информацию о структуре речевого сигнала с учетом нестационарных амплитуд, периодов и фаз гармоник, входящих в его состав. Данное свойство обусловлено тем, что пространство собственных векторов образует нестационарный базис, в который проецируется речевой сигнал. Однако повышение точности вычисления ЧОТ приводит к увеличению вычислительной сложности. Предлагаемая в статье модель сингулярного оценивания частоты основного тона позволяет оптимизировать временную обработку речевого сигнала за счет аппроксимации края сингулярного спектра, выделяя главные компоненты, образующие речевой сигнал для случая неизвестных априорных распределений амплитуд, периодов и начальных фаз гармоник. Далее рассматривается программная реализация модели и проводится сравнение с аналогами. В результате, в данной работе предлагается новый подход к оцениванию частоты основного тона речевого сигнала.

SUMMARY

Volf D.A., Meshcheryakov R. V. **Software Implementation of a Singular Meter of the Pitch Frequency of a Speech Signal.**

The article discusses a singular model of the speech signal, which allows considering speech production as a system of acoustic resonators in which the parameters are the eigenvalues and eigenvectors containing the information about the structure of the speech signal, taking into account time-dependent amplitudes, periods and phases of the harmonics included in its composition. This property is determined by the fact that the space of eigenvectors forms a transient basis, in which the speech signal is projected. Improving the accuracy of calculating the pitch frequency leads to higher computational complexity. The proposed in the article singular model of estimating the pitch frequency allows optimizing the time processing of the natural speech signal by approximating the edges of the singular spectrum, highlighting the main components that form a voice signal for the case of unknown a priori distributions of amplitudes, periods and the initial phases of the harmonics. In addition, software implementation of the model and a comparison with analogues are considered. As a result, this paper proposes a new approach to estimating the pitch frequency of the speech signal.

УДК 621.391.1

В.И. САЛУХОВ, В.С. СОЛДАТЕНКО

СТРУКТУРНО-ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ И МЕТОДИКА РЕШЕНИЯ ЗАДАЧИ ОБОСНОВАНИЯ МОДЕРНИЗАЦИИ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ

Салухов В.И., Солдатенко В.С. **Структурно-функциональная модель и методика решения задачи обоснования модернизации телекоммуникационных систем.**
Аннотация. В статье рассматривается подход к обоснованию модернизации сети связи по внедрению новых коммуникационных услуг. Критерием оптимальности доработок сети является экономическая эффективность. Анализируются проблемы внедрения новых коммуникационных услуг. Приводятся структурно-функциональная модель модернизации сети связи и модель принятия решений по выбору оптимальных по заданному критерию доработок каждой системы сети связи. Результаты моделирования иллюстрируются расчетным примером.
Ключевые слова: коммуникационная услуга, модернизация, критерий оптимальности, принятие решения.

Salukhov V.I., Soldatenko V.S. **Structurally Functional Model and Technique to Solve the Problem of Justification of Telecommunication Systems Modernization.**
Abstract. In the article, we consider an approach to justification of communication network modernization on introduction of new communication services. Optimality criterion of network completions is economic efficiency. Problems of introduction of new communication services are analyzed. The structurally functional model of communication network modernization and a model of decision-making at the choice of optimum by the set criterion of completions of each communication network system are given. Results of modeling are illustrated by a settlement example.
Keywords: communication service, modernizaition, optimality criterion, decision making.

1. Введение. В статье рассматривается подход к принятию решения по обоснованию варианта модернизации сети связи для внедрения перспективных коммуникационных услуг (ПКУ), оптимального по критерию экономической эффективности. Анализируются проблемы развития коммуникационных услуг, приводятся модель модернизации сети связи и модель принятия решений по выбору оптимального по установленному критерию варианта модернизации для каждой составной части сети связи, методика решения указанной задачи, а также результаты расчетов в соответствии с предложенными моделями.

2. Краткий анализ проблем развития коммуникационных услуг. В последние годы отрасль связи и телекоммуникаций является одной из наиболее динамично развивающихся в составе российской экономики. Об этом говорят темпы роста отрасли, которые несколько лет подряд существенно превышают рост валового внутреннего продукта (ВВП) России в целом в эти же интервалы времени [1]. В

сфере связи и телекоммуникаций сосредоточено более 70 % наукоемких производств и значительная доля капиталовложений. Создана огромная по своим масштабам инфраструктура, позволяющая оказывать большой объем разнообразных услуг связи [2, 3]. Однако в последнее время отмечается такая особенность обеспечения услугами связи, как опережающее моральное старение телекоммуникационных систем. При этом фактор физического износа оборудования оказывает гораздо меньшее воздействие на результативность функционирования действующих сетей связи. Это обусловлено появлением новых типов услуг связи. К ним относятся, например, инфокоммуникационные услуги [4, 5], услуги в рамках Концепции Интернета вещей (Internet of Things, IoT) и некоторые другие [6, 7, 8]. Предполагается, что активное внедрение перспективных коммуникационных услуг является в настоящее время одним из основных факторов, определяющих экономический рост развитых государств, развивающихся стран и стран с переходной экономикой.

Доступность перспективных коммуникационных услуг для пользователей вне зависимости от способов доступа может быть обеспечена лишь при конвергенции («взаимопроникновении») существующих сетей [9, 10]. Однако конвергенция сетей, обусловленная необходимостью одновременной передачи разными категориями пользователей различных видов информации в реальном времени, породила две глобальные технические проблемы [10]:

– необходимость поддержки большого разнообразия систем сигнализации, базирующихся на технологиях TDM, ATM, IP, MPLS и др.;

– необходимость достижения «конвергенции услуг связи» (наряду с «конвергенцией сетей»), с помощью которой осуществляется ввод новых коммуникационных услуг с универсальным доступом из сетей связи общего пользования (ССОП), сетей ISDN, интеллектуальной сети (IN), сети IP и др.

Решение этих задач осуществляется с помощью аппаратно-программных средств нового типа: «программных коммутаторов (Softswitch)», медиашлюзов (MGW) и ряда других. В целом, для создания возможностей использования ПКУ сети связи должны соответствовать ряду требований, основными из которых являются мультисервисность, широкополосность, мультимедийность, интеллектуальность, инвариантность доступа и многооператорность [9, 10]. Существующие ССОП, в которых используются коммутация каналов (для телефонных сетей общего

пользования – ТфОП) и коммутация пакетов (для сетей передачи данных – СПД), в настоящее время не отвечают перечисленным выше требованиям. Именно поэтому в отрасли телекоммуникаций интенсивно прорабатываются различные аспекты перехода к сети связи следующего поколения (ССП, Next Generation Network, NGN) [9, 10]. В основу концепции построения указанной телекоммуникационной системы положена идея о создании универсальной сети. Эта сеть должна позволять переносить любые виды информации, такие как речь, видео, аудио, графику и т. д., а также обеспечивать возможность предоставления самого широкого спектра перспективных коммуникационных услуг [10]. Однако быстрая замена действующих сетей связи перспективными мультисервисными системами во многих случаях является неприемлемой по финансовым, технологическим, социальным и иным соображениям. Более реалистическим направлением создания условий для предоставления перспективных коммуникационных услуг является поэтапная доработка или модернизация существующих сетей связи. Однако и при данном варианте развития телекоммуникационных систем необходимы достаточно большие финансовые, временные и трудовые ресурсы. Следует также учитывать, что имеется значительная неопределенность в прогнозах коммерческой успешности перехода к ПКУ и привлечения достаточно большого числа пользователей указанных услуг в ближайшие годы. Поэтому существует острая необходимость обоснования модернизации существующих сетей связи для обеспечения внедрения новых коммуникационных услуг в условиях ограниченных ресурсов. В настоящей статье рассматривается один из возможных путей решения указанной задачи.

3. Содержательная постановка задачи обоснования модернизации сети связи. Будем представлять модернизацию сети связи как совокупность доработок систем сети, которые обеспечивают внедрение соответствующих перспективных коммуникационных услуг. В данном случае функции сети связи (состав ПКУ) в целом определяются структурой составляющих ее систем и их возможностями. Поэтому решение указанной задачи целесообразно осуществлять с использованием класса структурно-функциональных

моделей [11]. Представим модернизацию сети связи в виде следующей иерархической модели (схема на рисунке 1)

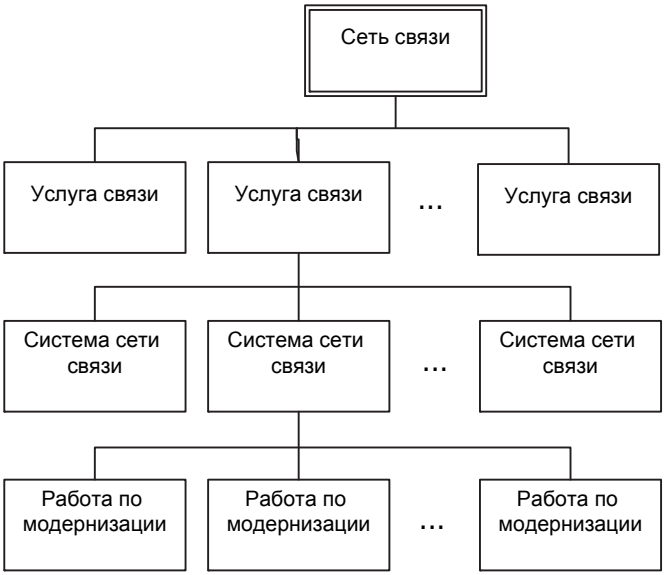


Рис. 1. Структурно-функциональная модель модернизации телекоммуникационной системы

Каждая работа по совершенствованию системы сети связи требует затрат соответствующих ресурсов. При этом результативность функционирования этих объектов в общем случае также может быть различной и зависеть от вида проводимой доработки. Следует также принимать во внимание ограниченность ресурсов каждого типа. Таким образом, для решения задачи обоснования модернизации сети связи необходимы следующие исходные данные:

– состав внедряемых перспективных коммуникационных услуг;

– состав модернизируемых систем сети связи, обеспечивающих реализацию соответствующих услуг связи;

– виды возможных доработок каждой из систем сети связи;

– затраты ресурсов на проведение доработок систем сети связи;

– предельно допустимые значения ресурсов каждого типа;

– предполагаемые результаты от совершенствования соответствующих систем связи.

Будем исходить из того, что выполнение работ по модернизации приводит к положительному изменению показателя результативности функционирования каждой из систем и сети связи в целом. Назовем это изменение откликом на выполнение работ по совершенствованию телекоммуникационной системы. С учетом коммерческого характера оказания услуг связи [2, 11] показателем результативности модернизируемой сети связи может выступать возможный приток денежных средств, который зависит от количества возможных пользователей перспективными коммуникационными услугами и размера платы за соответствующую услугу [12 – 15]. Из-за различий в количественных показателях исходных данных величина отклика на модернизацию сети связи также будет различной. Важнейшим необходимым условием осуществления варианта модернизации выступает соответствие предстоящих расходов ресурсов на модернизацию их допустимым значениям. Это условие предполагает, что доработки некоторых систем (внедрение соответствующих возможных ПКУ) могут оказаться нецелесообразными.

После того, как сформирован перечень систем сети связи, которые должны быть усовершенствованы при внедрении предполагаемых ПКУ, и состав возможных работ для каждой из указанных систем, необходимо осуществить моделирование процесса выбора варианта модернизации, наиболее приемлемого по величине отклика сети связи. Для этого требуется сформировать множество допустимых вариантов модернизации существующей сети связи и осуществить выбор из этого множества варианта, оптимального по критерию экономической результативности. Применение данного подхода позволяет оценить целесообразность и состав доработок отдельных систем с учетом имеющихся возможностей для модернизации функционирующей сети связи в целом и прогнозируемых предпочтений ее пользователей.

Следует отметить, что в настоящее время имеется достаточно большая номенклатура однотипных средств, входящих в состав систем сети связи. Это приводит к значительному количеству вариантов их использования при модернизации сети. Очевидно, что количество вариантов по модернизации сети связи стремительно возрастает при увеличении каждого из компонентов модели, представленной на схеме рисунка 1. Задача оценивания результата для того или иного варианта совершенствования оборудования сети связи имеет высокую сложность из-за необходимости учета большого числа факторов, определяющих этот результат. Поэтому задача выбора наилучшего в некотором смысле варианта модернизации сети связи является актуальной для

современного этапа развития теории и практики эксплуатации телекоммуникационных систем. Рассмотрим последовательность этапов предлагаемой методики решения этой задачи.

1. Построение модели совершенствования отдельных систем сети связи, обеспечивающих оказание той или иной ПКУ.

2. Формирование частных показателей результативности доработок для каждой отдельной системы сети связи.

3. Формирование общего критерия оптимальности модернизации всей совокупности модернизируемых средств всех систем сети связи.

4. Формирование множества ограничений, влияющих на процесс модернизации сети связи. Это могут быть ограничения на использование ресурсов различных видов – финансовых, временных, трудовых и др.

5. Построение модели выбора варианта модернизации для всей совокупности систем сети связи, реализующих предполагаемые ПКУ.

6. Определение наилучшего в смысле выбранного критерия оптимальности варианта модернизации совокупности систем сети связи с учетом действующих ограничений на процесс модернизации.

Рассмотрим математическую постановку рассмотренной задачи.

4. Математическая модель оптимальной модернизации сети связи. Для упрощения необходимых математических соотношений будем полагать, что для внедрения отдельно взятой перспективной коммуникационной услуги необходимо совершенствование только одной соответствующей системы сети связи. Обозначим через N общее количество предполагаемых для внедрения ПКУ. Таким образом, наибольшее количество систем сети связи, которые могут быть модернизированы, также равно N. Следует учесть, что с учетом сделанных предположений при проведении модернизации сети связи количество $N^{(\phi)}$ фактически внедренных ПКУ может отличаться от N. При этом должно соблюдаться условие:

$$N^{(\phi)} \le N \qquad (1)$$

Установим, что для каждой i-й системы возможно некоторое количество K_i ($i = 1(1)N$) видов работ по ее совершенствованию. Обозначим через k_i порядковый номер возможной доработки i-й системы ($k_i = 1(1)K_i$, $i = 1(1)N$). Каждая доработка соответствует возможным действиям по совершенствованию системы сети связи, например, замене отдельных аппаратных или программных средств

системы на новые, необходимое тестирование и др. Выбранный для реализации номер доработки i-й системы будем обозначать k_i^*. При совершенствовании программных средств связи, как правило, требуется осуществлять как автономное, так и общее тестирование программного обеспечения всей системы связи. При моделировании установим, что эти процедуры входят в состав проводимых доработок.

Показателями совершенствования системы сети связи являются затраты ресурсов различной природы на реализацию доработки, а также получаемый эффект от использования соответствующей ПКУ. Ресурсами могут быть выделяемые для модернизации объем финансирования, допустимая длительность времени, количество специалистов и другие. Пусть количество видов принимаемых во внимание ресурсов равно R. Обозначим затрачиваемые при вариантах модернизации i-й системы ресурсы через $g_{ir}(k_i)$ ($k_i = 1(1)K_i$, $i = 1(1)N$, $r = 1(1)R$), расчетные затраты ресурсов через $G_r^{(расч)}(\{k_i^*\}, i \in 1(1)N^{(факт)})$, а допустимые значения каждого вида ресурса – через $G_r^{(доп)}$ ($r = 1(1)R$). Расчетное соотношение для $G_r^{(расч)}(\{k_i^*\}, i \in 1(1)N^{(факт)})$ имеет вид:

$$G_r^{(расч)}(\{k_i^*\}, i \in 1(1)N^{(факт)}) = \sum_{i \in \{1(1)N^{(\phi)}\}} g_{ir}(k_i^*), \quad r = 1(1)R. \tag{2}$$

Соотношение (2) определяет, что принимаются во внимание затраты ресурсов только на системы, подвергшиеся доработке. При этом для каждого r-го вида ресурса должны выполняться следующие условия:

$$G_r^{(расч)} \le G_r^{(доп)}, \quad r = 1(1)R. \tag{3}$$

Выходной эффект от внедрения i-й услуги в результате проведения k_i-го вида доработки i-й системы представляет собой приток $C_i(k_i)$ ($k_i = 1(1)K_i$, $i = 1(1)N$) финансовых средств за установленный интервал времени. Установим, что приток $C_i(k_i)$ ($k_i = 1(1)K_i$, $i = 1(1)N$) определяется количеством $L_i(k_i)$ пользователей i-й услугой и стоимостью $c_i(k_i)$ указанной услуги ($k_i = 1(1)K_i$, $i = 1(1)N$). В таком случае расчетное соотношение для притока по каждой возможной доработке i-й системы сети связи имеет следующий вид:

$$C_i(k_i) = c_i(k_i) \cdot L_i(k_i), \quad k_i = 1(1)K_i, \; i = 1(1)N. \tag{4}$$

Общий выходной эффект C_M от внедрения $N^{(\phi)}$ ПКУ определяется соотношением:

$$C_M = \sum_{i \in \{1(1)N^{(\phi)}\}} C_i(k_i^*) \,. \qquad (5)$$

Соотношение (5) показывает, что принимаются во внимание притоки только от ПКУ, реализуемых системами, прошедшими совершенствование. Таким образом, соотношения (2) – (5) определяют модель модернизации сети связи при внедрении перспективных коммуникационных услуг. Перейдем теперь к рассмотрению модели выбора оптимального по экономическому критерию варианта модернизации сети связи в условиях имеющихся ограничений по ресурсам. Для построения указанной модели применим математический аппарат теории принятия решений [16]. Это предполагает использование методики, содержащей следующие этапы:

1. Формирование целевой функции расчетной процедуры, которая характеризует результаты модернизации по выбранному критерию. Значения указанной целевой функции должны зависеть от номера k_i доработки каждой из N систем, используемых для внедрения новых коммуникационных услуг сети связи.
2. Определение множества допустимых вариантов модернизации.
3. Использование расчетной процедуры оптимизации, соответствующей свойствам целевой функции и ограничений задачи.

Дадим обобщенное описание введенных исходных данных в форме матрицы, содержащейся в таблице 1.

Таблица 1. Описание доработок систем сети связи

Система сети связи	Номер работы	Затраты ресурса 1-го типа	Затраты ресурса 2-го типа	…	Затраты ресурса R-го типа	Число польз-лей $L_i(k_i)$	Размер платы за ПКУ $c_i(k_i)$
Система 1	$k_1=1$	$g_{11}(1)$	$g_{12}(1)$	…	$g_{1R}(1)$	$L_1(1)$	$c_1(1)$
	$k_1=2$	$g_{11}(2)$	$g_{12}(2)$	…	$g_{1R}(2)$	$L_1(2)$	$c_1(2)$
	…	…	…	…	…	…	…
	$k_1=K_1$	$g_{11}(K_1)$	$g_{12}(K_1)$	…	$g_{1R}(K_1)$	$L_1(K_1)$	$c_1(K_1)$
…	…	…	…	…	…	…	…
Система N	$k_N=1$	$g_{N1}(1)$	$g_{N2}(1)$	…	$g_{NR}(1)$	$L_N(1)$	$c_N(1)$
	$k_N=2$	$g_{N1}(2)$	$g_{N2}(2)$	…	$g_{NR}(2)$	$L_N(2)$	$c_N(2)$
	…	…	…	…	…	…	…
	$k_N=K_N$	$g_{N1}(K_N)$	$g_{N2}(K_N)$		$g_{NR}(K_N)$	$L_N(K_N)$	$c_N(K_N)$

Множество допустимых вариантов модернизации определяется ограничениями вида (3), а также свойствами используемых в задаче переменных. В данной работе критерием оптимальности модернизации сети связи выступает условие достижения максимального значения притока C_M от всей совокупности $N^{(\phi)}$ внедренных ПКУ. Значение этой функции определяется с помощью соотношения (5).

Поставим в соответствие каждой переменной k_i булеву переменную $x(k_i)$, которая является индикатором выбора соответствующей работы для i-й системы ($k_i = 1(1)K_i$, $i = 1(1)N$). При этом $x(k_i) = 1$, если выбрана работа k_i, и $x(k_i) = 0$ в противном случае ($k_i = 1(1)K_i$, $i = 1(1)N$). Таким образом, переменные $x(k_i)$ являются неизвестными переменными задачи определения оптимального по критерию максимальной экономической эффективности варианта модернизации сети связи. Ограничения в указанной задаче соответствуют условиям выполнения требований непревышения необходимых затрат ресурсов каждого типа их установленных предельных значений, выбора не более одного варианта доработок каждой из N анализируемых систем сети связи и булевого характера неизвестных переменных задачи. Таким образом, математическая модель рассматриваемой задачи оптимизации определяется следующими соотношениями:

$$C_M(\mathbf{x}) \to \max_{\mathbf{x} \in \Delta_\beta}, \tag{6}$$

при этом:

$$C_M(\mathbf{x}) = \sum_{i=1}^{N} C_i(\mathbf{x}), \tag{7}$$

где

$$C_i(\mathbf{x}) = \sum_{k_i=1}^{K_i} c_i(k_i) \cdot L_i(k_i) \cdot x(k_i), \; i = 1(1)N, \tag{8}$$

а также:

$$\Delta_\beta = \{\Delta \Big| G_r^{(расч)}(\mathbf{x}) \le G_r^{(доп)}; \\ \sum_{k_i=1}^{K_i} x(k_i) \le 1; \; x(k_i) \in \{0,1\} \; (k_i = 1(1)K_i, i = 1(1)N)\}, \tag{9}$$

при этом:

$$G_r^{(расч)}(\mathbf{x}) = \sum_{i=1}^{N} \sum_{k_i=1}^{K_i} g_{ir}(k_i) \cdot x(k_i), \quad r = 1(1)R, \tag{10}$$

$$\begin{cases} x(k_i) = 1, \text{ при } k_i = k_i^*; \\ x(k_i) = 0, \text{ при } k_i \neq k_i^*. \end{cases}, \quad k_i = 1(1)K_i, \; i = 1(1)N. \tag{11}$$

Соотношение (9) характеризует формирование множества допустимых вариантов модернизации сети, определяя необходимость выполнения следующих условий: расчетные затраты ресурсов на модернизацию не должны превышать их допустимых значений; проведение на i -й системе только одной из предусмотренных для нее доработок, либо отсутствие доработок на системе; а также возможные значения булевых переменных задачи. Соотношение (10) определяет значения расчетных затрат ресурсов на модернизацию сети связи. Соотношение (11) характеризует условия определения значений булевых переменных $x(k_i)$ ($k_i = 1(1)K_i$, $i = 1(1)N$). Выберем теперь класс методов решения задачи оптимизации (6) – (11). Поскольку переменные $x(k_i)$ ($k_i = 1(1)K_i$, $i = 1(1)N$) являются булевыми, то процедура решения задачи оптимизации должна соответствовать данной особенности модели. Сформированная модель имеет конечную мощность множества допустимых альтернатив Δ_β. Это объясняется конечностью количества систем сети связи и конечностью количества возможных доработок этих объектов. При этом мощность пустого множества Δ_β равна 0, что соответствует случаю полного отказа от проведения доработок каких-либо систем сети связи. Кроме этого, целевая функция и ограничения задачи оптимизации являются линейными. Указанные особенности позволяют сделать вывод о том, что для данной задачи можно получить оптимальное решение [17]. Для иллюстрации применения представленной модели рассмотрим расчетный пример.

5. Расчетный пример. Имеется сеть связи. Предполагается внедрение шести перспективных коммуникационных услуг. Для реализации каждой из указанных услуг необходима только одна система сети связи. Возможные варианты по доработкам (в ходе модернизации) указанных систем характеризуются расходом необходимых ресурсов. Принимаются во внимание два вида ресурсов: финансовые расходы и затраты времени на проведение необходимых работ.

Предполагается, что работы по модернизации систем проводятся параллельно. Следовательно, длительность модернизации сети связи определяется продолжительностью доработки системы с самой большой длительностью. На осуществление модернизации установлены предельный размер финансирования и предельное время модернизации. Выходной эффект от варианта модернизации каждой из систем сети связи определяется количеством пользователей, которым по техническим возможностям могут быть предоставлены внедряемые услуги. Определена также прогнозная стоимость оказания каждой ПКУ пользователю. Необходимо определить, какие системы и по какому варианту необходимо дорабатывать в ходе модернизации сети связи, чтобы добиться максимального экономического эффекта.

Числовые исходные данные по доработкам систем сети связи приведены в таблице 2.

Таблица 2. Исходные данные по доработкам систем сети связи

Номер системы (i)	Номер доработки (k_i)	Затраты финансов на доработку $g_{i1}(k_i)$ (усл. ед.)	Затраты времени на доработку $g_{i2}(k_i)$ (ед. вр.)	Возможное количество пользователей ($L_i(k_i)$)	Размер платы за ПКУ $c_i(k_i)$ (усл. ед.)
1	1	3000	400	200	10
	2	4000	600	300	10
	3	6000	800	600	10
2	1	1000	450	700	40
	2	2500	650	900	40
3	1	2800	700	420	50
	2	3200	800	560	50
	3	4100	1000	810	50
4	1	1200	620	580	20
	2	2300	780	640	20
	3	3800	1200	820	20
5	1	1600	390	330	30
	2	2200	440	420	30
	3	3400	680	590	30
	4	5600	960	920	30
6	1	5200	860	910	70
	2	6800	1100	1400	70

Значения k_i ($k_i = 1(1)K_i$, $i = 1(1)N$) приведены в столбце 2 табл. 2. При этом $K_1 = 3$; $K_2 = 2$; $K_3 = 3$; $K_4 = 3$; $K_5 = 4$; $K_6 = 2$. Аналогичным образом представлены другие исходные данные по видам доработок систем сети связи.

На проведение модернизации сети связи выделено 30 000 усл. ед. При этом длительность модернизации не может превышать 700 единиц времени (ед. вр.). Таким образом, в соответствии с введенными обозначениями можно записать следующее: количество возможных ПКУ $N = 6$; количество видов ресурсов $R = 2$, $G_1^{(don)} = 30000$ усл. ед.; $G_2^{(don)} = 700$ ед. вр.

*Необходимо***.** Определить, какие системы необходимо совершенствовать и с помощью каких доработок k_i^* ($i = 1(1)N$) при модернизации сети связи, чтобы получить наибольший приток от предоставления внедряемых ПКУ пользователям при выполнении требований по имеющимся ограничениям по ресурсам.

*Решение***.** Введем булевы неизвестные переменные задачи $\mathbf{x}$. В начале решения установим, что $x(k_i) = 0$ ($k_i = 1(1)K_i$, $i = 1(1)N$). Для исходных данных задачи общее число искомых переменных равно 17. Таким образом, решением задачи определения оптимального по критерию максимального конечного притока варианта модернизации систем сети связи является определение значений элементов вектора $\mathbf{x}$ индикаторов выбора соответствующей доработки.

Для решения задачи использован метод ветвей и границ. Программным средством для решения задачи выбрана надстройка MS-Excel «Поиск решения» [18]. На начальном этапе решения присвоим всем неизвестным нулевые значения. Получены следующие результаты.

Совершенствованию при модернизации сети связи подлежат первая, вторая, третья и пятая системы. Оптимальными видами доработок систем сети связи являются следующие: $k_1^* = 2$; $k_2^* = 2$; $k_3^* = 1$; $k_5^* = 3$. Это означает, что при внедрении перспективных коммуникационных услуг необходимо доработать первую систему по варианту 2, вторую систему – по варианту 2; третью систему – по варианту 1; пятую систему – по варианту 3. Внедрять четвертую и шестую коммуникационные услуги с доработками соответствующих систем сети связи является в данном случае нецелесообразным. Сводные результаты расчетов представлены в таблице 3.

Таблица 3. Результаты решения задачи обоснования оптимальной модернизации сети связи

Номер системы (i)	Номер (k_i)	Значение $x(k_i)$	Затраты финансов $g_{i1}(k_i)$ (усл. ед.)	Затраты времени $g_{i2}(k_i)$ (ед. вр.)	Возможное количество пользователей ($L_i(k_i)$)	Размер платы за ПКУ $c_i(k_i)$ (усл. ед.)	Приток C_i (усл. ед.)
1	0	0	0	0	0	0	3000
	2	1	4000	600	300	10	
	0	0	0	0	0	0	
2	0	0	0	0	0	0	36000
	2	1	2500	650	900	40	
3	1	1	2800	700	420	50	21000
	0	0	0	0	0	0	
	0	0	0	0	0	0	
4	1	0	0	0	0	0	0
	0	0	0	0	0	0	
	0	0	0	0	0	0	
5	0	0	0	0	0	0	17700
	0	0	0	0	0	0	
	3	1	3400	680	590	30	
	0	0	0	0	0	0	
6	0	0	0	0	0	0	0
	0	0	0	0	0	0	

При этом на модернизацию сети связи требуется потратить $G_1^{(расч)} = 12700$ усл. ед. Необходимые работы по модернизации должны быть выполнены за $G_2^{(расч)} = 700$ ед. вр. Достигаемый при модернизации экономический результат, характеризуемый значением притока за внедряемые при модернизации сети связи коммуникационные услуги, составляет $C_M = 77700$ усл. ед. Таким образом, решение расчетного примера завершено.

6. Заключение. В данной статье рассмотрено структурно-функциональное моделирование модернизации сети связи, которое позволяет обосновать состав типов доработок систем сети, оптимальный по критерию максимального притока денежных средств за пользование внедренными коммуникационными услугами. При этом модель оптимизации доработок представлена как задача булева математического программирования, решаемая методом ветвей и границ. Исследование доведено до инженерного решения на основе использования математического пакета MS-Excel. Выбор данного вычислительного средства упрощает практическое использование

разработанного структурно-функционального моделирования модернизации сети связи. Это обусловлено доступностью и распространенностью MS-Office, отсутствием необходимости дополнительного обучения использованию соответствующего математического пакета. Предложенная форма представления исходных данных, промежуточных результатов счета и результатов обоснования варианта модернизации сети связи достаточно наглядна и удобна для анализа.

Практическая значимость полученных результатов состоит в возможности их использования при проектировании сетей связи, а также в экспертных организациях отрасли связи, осуществляющих анализ обоснованности инвестиций в сфере телекоммуникаций.

Литература

1. Cnews. Издание о высоких технологиях. URL: cnews.ru (дата обращения: 25.11.2015).
2. Федеральный закон от 7 июля 2003 г. № 126-ФЗ «О связи». URL: docs.cntd.ru (дата обращения: 10.07.2015).
3. *Величко В.В. Субботин Е.А., Шувалов В.П., Ярославцев А.Ф.* Телекоммуникационные системы и сети. Современные технологии. Том 3. Мультисервисные сети // М.: Горячая линия-Телеком. 2005. 592 с.
4. *Линец Г.И., Фомин Л.А., Говорова С.В., Меденец В.В.* Построение мультисервисных сетей на основе функциональных преобразований трафика // Инфокоммуникационные технологии. 2014. Том 12. № 4. С. 40-45.
5. *Буданов А.Н., Дмитриев В.М.* Виртуальные интерфейсы для цифровых систем передачи мультисервисного трафика // Инфокоммуникационные технологии. 2013. Том 11. № 3. С. 27-30.
6. *Сарьян В.К. и др.* Прошлое, настоящее и будущее стандартизации интернета вещей // Труды научно-исследовательского института радио. 2014. № 1. С. 2–7.
7. *Титаренко Е.* Российские преграды для IoT. URL: www.comnews.ru. (дата обращения: 09.07.2015).
8. Yang L., Yang S.H., Plotnick L. How the internet of things technology enhances emergency response operations // Technological Forecasting and Social Change. 2013. vol. 80(9). pp. 1854–1867.
9. *Бутенко В.В., Назаренко А.П., Сарьян В.К.* и др. Проблемы, возникающие при внедрении новых технологий в инфокоммуникационном сообществе // Труды НИИР. 2011. № 1. С. 12–19.
10. *Кучерявый А.Е., Цуприков А.Л.* Сети связи следующего поколения // М.: ФГУП ЦНИИС. 2006. 278 с.
11. *Цвиркун А.Д.* Основы синтеза структуры сложных систем // М.: Наука. 1982. 200 с.
12. *Резникова Н.П.* Маркетинг в телекоммуникациях // М.: Эко-Трендз. 2002. 336 с.
13. Wallin S., Leijon V. Multi-Purpose Models for QoS monitoring // In 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW’07). IEEE Computer Society. 2007. pp. 900–905.
14. Carvalho de Gouveia F., Magadanz T. Quality of service in telecommunication network // Telecommunication Systems and Technologies. 2008. vol. 2. 21 p.

15. Amani N., Alipour E. Analysis of Performance and Quality Parameters for Service Level Agreement in Long Distance Calls Service // Proceedings of the International MultiConference of Engineers and Computer Scientists. Hong Kong. 2008. vol. 2. 4 p.
16. *Черноруцкий И.Г.* Методы принятия решений. Учебное пособие // СПб.: БХВ – Петербург. 2005. 408 с.
17. *Резников Б.А.* Методы и алгоритмы оптимизации на дискретных моделях сложных систем // Л.: ВИКИ. 1983. 250 с.
18. *Леоненков А.В.* Решение задач оптимизации в среде MS Excel // СПб.: БХВ-Петербург. 2005. 704 с.

References

1. Cnews. Izdanie o vysokih tehnologijah. [Cnews. The edition about high technologies]. URL: cnews.ru (accessed 25.11.2015). (In Russ.).
2. Federal'nyj zakon ot 7 ijulja 2003 g. № 126-FZ. [«About communication»]. Available at: docs.cntd.ru (accessed 10.07.2015). (In Russ.).
3. Velichko V.V. Subbotin E.A., Shuvalov V.P., Jaroslavcev A.F. *Telekommunikacionnye sistemy i seti. Sovremennye tehnologii. Tom 3. Mul'tiservisnye seti.* [Telecommunication systems and networks. Modern technologies. Volume 3. Multiservice networks]. M.: Gorjachaja linija-Telekomю. 2005. 592 p. (In Russ.).
4. Linec G.I., Fomin L.A., Govorova S.V., Medenec V.V. [Creation of multiservice networks on the basis of functional transformations of a traffic]. *Infokommunikacionnye tehnologii – Infocommunication technologies*. 2014. vol. 12. no. 4. pp. 40–45. (In Russ.).
5. Budanov A.N., Dmitriev V.M. [Virtual interfaces for digital systems of transfer of a multiservice traffic]. *Infokommunikacionnye tehnologii – Infocommunication technologies*. 2013. vol. 11. no. 3. pp. 27–30. (In Russ.).
6. Sar'jan V.K., et al. [Last, real and future standardization of the Internet of things]. *Trudy nauchno-issledovatel'skogo instituta radio – Works of research institute of radio*. 2014. vol. 1. pp. 2–7. (In Russ.).
7. Titarenko E. Rossijskie pregrady dlja IoT [Russian barriers to IoT]. Available at: www.comnews.ru (accessed 09.07.2015). (In Russ.).
8. Yang L., Yang S.H., Plotnick L. How the internet of things technology enhances emergency response operations. *Technological Forecasting and Social Change*. 2013. vol. 80(9). pp. 1854–1867.
9. Butenko V.V., Nazarenko A.P., Sar'jan V.K., et al. [The problems arising at introduction of new technologies in infocommunication community]. *Trudy NIIR – Proceedings of NIIR*. 2011. vol. 1. pp. 12–19. (In Russ.).
10. Kucherjavyj A.E., Cuprikov A.L. *Seti svjazi sledujushhego pokolenija*. [Communication networks of the next generation]. M.: FGUP CNIIS. 2006. 278 p. (In Russ.).
11. Cvirkun A.D. *Osnovy sinteza struktury slozhnyh sistem* [Bases of synthesis of structure of difficult systems]. M.: Nauka. 1982. 200 p. (In Russ.).
12. Reznikova N.P. *Marketing v telekommunikacijah* [Marketing in telecommunications]. M.: Jeko-Trend. 2002. 336 p. (In Russ.).
13. Wallin S., Leijon V. Multi-Purpose Models for QoS monitoring. In 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07). IEEE Computer Society. 2007. pp. 900–905.
14. Carvalho de Gouveia F., Magadanz T. Quality of service in telecommunication network. *Telecommunication Systems and Technologies*. 2008. vol. 2. 21 p.
15. Amani N., Alipour E. Analysis of Performance and Quality Parameters for Service Level Agreement in Long Distance Calls Service. Proceedings of the International MultiConference of Engineers and Computer Scientists. Hong Kong. 2008. vol. 2. 4 p.

16. Chernoruckij I.G. *Metody prinjatija reshenij. Uchebnoe posobie.* [Decision-making methods. Tutorial]. SPb.: BHV-Peterburg. 2005. 408 p. (In Russ.).
17. Reznikov B.A. *Metody i algoritmy optimizacii na diskretnyh modeljah slozhnyh sistem*. [Methods and algorithms of optimization on discrete models of difficult systems]. L.: VIKI. 1983. 250 p. (In Russ.).
18. Leonenkov A.V. *Reshenie zadach optimizacii v srede MS Excel*. [The solution of problems of optimization in the environment of MS Excel]. SPb.: BHV-Peterburg. 2005. 704 p. (In Russ.).

Салухов Владимир Иванович — к-т техн. наук, доцент, руководитель исследовательской группы информационных технологий в образовании, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: информационные технологии в образовании, управление жизненным циклом инфотелекоммуникационных систем, анализ и разработка систем поддержки принятия решений на базе современных информационных технологий. Число научных публикаций — 50. vsaluhov@bk.ru; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7(812)3280382.

Salukhov Vladimir Ivanovich — Ph.D., associate professor, head of research group of information technologies in education, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: research and information technologies in educaion, lifecycle managment infocommunication systems, analysis and development of support systems and decision making on the basis of modern information yechnologies. The number of publications — 50. vsaluhov@bk.ru; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)3280382.

Солдатенко Владимир Стальевич — к-т техн. наук, доцент, Член-корреспондент Российской муниципальной академии, доцент кафедры метрологии, Военно-космическая академия имени А.Ф. Можайского (ВКА им. А.Ф.Можайского), доцент кафедры гражданского строительства и прикладной экологии, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ). Область научных интересов: информационные технологии в образовании, управление жизненным циклом инфотелекоммуникационных систем, анализ и разработка систем поддержки и принятия решений на базе современных информационных технологий. Число научных публикаций — 62. soldatenko_vs@mail.ru; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7(812)328-0103, Факс: +7(812) 328-4450.

Soldatenko Vladimir Stalyevich — Ph.D., associate professor, associate professor of civil engineering and applied ecology department, St. Petersburg polytechnical university of Peter the Great, associate professor of metrology dapartment, Mozhaisky Military Space Academy. Research interests: information technologies in education, lifecycle management infocommunication systems, analysis and development of support systems and decision-making on the basis of modern information technologies. The number of publications — 62. soldatenko_vs@mail.ru; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)328-0103, Fax: +7(812) 328-4450.

Поддержка исследований. Работа выполнена при финансовой поддержке Программы фундаментальных научных исследований ОНИТ РАН (проект № 2.11).

Acknowledgements. This research is supported by Program of fundamental scientific research ONIT Russian Academies of Sciences (grant 2.11).

РЕФЕРАТ

Салухов В.И., Солдатенко В.С. **Структурно-функциональная модель и методика решения задачи обоснования модернизации телекоммуникационных систем.**

Объектом исследования в статье является управление эволюционным развитием телекоммуникационных систем в направлении внедрения новых коммуникационных услуг связи.

Цель статьи – формирование подхода для обоснования необходимости доработок систем сети связи и выбора их вариантов, составляющих модернизацию сети при внедрении новых коммуникационных услуг связи, на основе учета перспективного спроса на указанные услуги и имеющихся ресурсов.

В представленных материалах обосновано применение структурно-функционального моделирования к описанию модернизации телекоммуникационной системы. Это позволяет учитывать влияние доработок систем сети связи на результат использования каждой новой коммуникационной услуги с учетом потребностей пользователей. Такой подход создает возможность формирования вариантов модернизации сети связи, подлежащих дальнейшему анализу. Разработаны предложения по получению оптимальной совокупности доработок систем сети связи по критерию максимальной экономической эффективности на основе математического целочисленного программирования.

Практическое применение полученных в статье результатов позволит создать научный аппарат оценивания обоснованности проектов развития сетей связи в аспекте их модернизации с ориентированием на спрос в новых коммуникационных услугах.

SUMMARY

Salukhov V.I., Soldatenko V.S. **Structurally Functional Model and Technique to Solve the Problem of Justification of Telecommunication Systems Modernization.**

The object of the research is management of the evolutionary development of telecommunication systems towards the introduction of new communication services.

The purpose of the article is to form an approach for justification of a need for communication network systems completions and a choice of their options making modernization of a network when implementing communication services on the basis of the accounting of perspective demand for the specified services and the available resources.

In the presented materials, application of structurally functional modeling to the description of modernization of telecommunication system is proved. It allows one to consider the influence of communication network systems completions on the result of the use of each new communication service taking into account the needs of users. Such an approach makes it possible to form the communication network modernization options, which are subject to the further analysis. Offers on obtaining optimum set of communication network systems completions on criterion of the maximum economic efficiency on the basis of mathematical integer programming are developed.

Practical application of the results received, presented in the article, will allow one to create the scientific apparatus to estimate validity of development projects of communication networks in aspect of their modernization with orientation on the demand for new communication services.

УДК 004.056.53

А.В. Козачок, М.В. Бочков, Р.Р. Фаткиева, Л.М. Туан

АНАЛИТИЧЕСКАЯ МОДЕЛЬ ЗАЩИТЫ ФАЙЛОВ ДОКУМЕНТАЛЬНЫХ ФОРМАТОВ ОТ НЕСАНКЦИОНИРОВАННОГО ДОСТУПА

Козачок А.В., Бочков М.В., Фаткиева Р.Р., Туан Л.М. **Аналитическая модель защиты файлов документальных форматов от несанкционированного доступа.**
Аннотация. В статье представлено описание метода защиты от несанкционированного доступа, основанного на применении процедуры неразличимой обфускации. Обосновано применение неразличимой обфускации для решения задачи защиты от несанкционированного доступа. Предложена математическая модель неразличимой обфускации программного кода, положенная в основу метода защиты от несанкционированного доступа к файлам документальных форматов.
Ключевые слова: обфускация, булева функция, несанкционированный доступ, защита.

Kozachok A.V., Bochkov M.V., Fatkieva R.R., Tuan L.M . **Analytical Model for Protecting Documentary File Formats from Unauthorized Access.**
Abstract. The article describes the method of documentary file formats protection from unauthorized access based on indistinguishable program code obfuscation. The application of indistinguishable obfuscation to solve the problem of unauthorized access protection is substantiated. Mathematical model of indistinguishable program code obfuscation underlying the method of documentary file formats protection from unauthorized access is proposed.
Keywords: obfuscation, Boolean function, unauthorized access, protection.

1. Введение. В последние годы особое внимание уделяется вопросам обеспечения безопасности инфокоммуникационных систем. Главным образом, эти исследования касаются формальной верификации свойств безопасности. Основной целью при этом является разработка формальной математической модели свойств безопасности в системе, а также верификация этой модели с помощью математических доказательств.

Для обеспечения безопасности инфокоммуникационных систем в течение многих лет акцент делался на обеспечение контроля доступа субъектов к объектам доступа. Невозможность полностью устранить утечку информации при этом подходе постепенно привела специалистов в области обеспечения безопасности к пониманию важности исследований информационных потоков в системе. Было предложено большое количество формальных моделей поведения системы и определения информационных потоков.

Целью проводимого исследования является построение комплекса моделей и алгоритмов процесса контролируемого разграничения доступа к файлам документальных форматов, позволяющего осуществить защиту от несанкционированного доступа к информации за счет применения неразличимой обфускации программного кода [1].

Исходя из вышеизложенного, в статье предлагается аналитическая модель защиты от несанкционированного доступа к файлам документальных форматов, отличающаяся применением процедуры неразличимой обфускации программного кода. При этом субъектами доступа в модели выступают пользователи, идентифицируемые учетными записями, а объектами являются файлы документальных форматов. Правила разграничения доступа субъектов к объектам задаются в виде матрицы полномочий, учитывающей метки конфиденциальности.

Предлагаемая модель позволяет хранить файлы документальных форматов в унифицированном виде и обеспечивает единый метод доступа к ним. Для безопасного хранения данных используется формат защищенного контейнера, в котором данные хранятся в обфусцированном виде. Контейнер представляет собой исполняемый файл, в который инкапсулируются файлы документальных форматов, обладающий рядом заданных свойств и функций, позволяющих однозначно идентифицировать пользователя, разграничивать доступ к данным, обеспечивать защиту конфиденциальности внедренного документа. Формат контейнера обеспечивает его безопасное хранение и передачу по сети [2].

2. Обзор исследований в области обфускации программного кода. Обфускацией программы называется всякое ее преобразование, которое сохраняет вычисляемую программой функцию, но при этом придает программе такую форму, что извлечение из текста программного кода ключевой информации об алгоритмах и структурах данных, реализованных в этой программе, становится трудоемкой задачей [3].

Обфусцированной программой называется программа, которая после применения обфусцирующих преобразований, на всех допустимых для исходной программы входных данных выдает тот же самый результат, что и оригинальная программа, но более трудна для анализа, понимания и модификации [2].

В настоящее время исследования в области обфускации программного кода проводятся по двум направлениям [3]:

– системное программирование;

– математическая криптография.

С позиции системного программирования обфускация программы может использоваться для защиты авторских прав на программное обеспечение, для предотвращения реверс инжиниринга программ, для создания и защиты водяных знаков, обеспечения безопасности мобильных агентов в информационных сетях, для проведения безопасного поиска в потоках данных и защиты баз данных. Однако существенным недостатком данного подхода является отсутствие обоснования гарантированной стойкости. В случае применения методов динамического анализа программ и привлечения квалифицированных экспертов

в области системного программирования стойкости существующих средств обфускации программ оказывается недостаточно.

С позиции математической криптографии разработка эффективных алгоритмов позволит решить целый ряд серьезных вопросов, например, с ее помощью можно преобразовать криптосистему с секретным ключом к криптосистеме с открытым ключом, проводить вычисления над зашифрованными данными, реализовывать системы функционального шифрования, доверенные схемы перешифрования и электронно-цифровой подписи, создавать верифицируемые системы тайного голосования и схемы двойственного шифрования.

Для построения эффективного метода защиты файлов документальных форматов, внедренных в защищенный контейнер, предлагается использовать математический аппарат неразличимой обфускации программного кода, активно развивающийся в настоящее время в рамках направления математической криптографии [4]. Исследования в области неразличимой обфускации, проводимые в настоящее время, как российскими учеными (Варнавский Н.П., Захаров В.А., Кузюрин Н.Н.), так и зарубежными (S. Garg, C. Gentry, S. Halevi, B. Barak, J.S. Coron, T. Lepoint, M. Tibouchi) базируются на возможности обфускации булевых функций. Процедуру проверки прав доступа пользователя к документу, внедренному в контейнер, можно рассматривать как точечную функцию, поскольку результатом ее выполнения является значение из множества $\{0,1\}$, поэтому применение неразличимой обфускации для защиты данной проверки является также корректным. Данная работа посвящена модификации существующих подходов к осуществлению неразличимой обфускации с целью устранения ряда ограничений, обусловленных применяемыми механизмами, моделями и алгоритмами [5, 6].

3. Описание аналитической модели защиты файлов документальных форматов от несанкционированного доступа. Для описания предлагаемой математической модели защиты от несанкционированного доступа к файлам документальных форматов, отличающейся применением процедуры неразличимой обфускации программного кода, необходимо ввести ряд обозначений:

iO – обфускатор неразличимости (неразличимый обфускатор);

$f(x)$ – булева функция, принимающая на вход вектор x длины n;

$C(x)$ – булева схема, принимающая на вход вектор x длины n;

$\mathbb{N}$ – множество натуральных чисел;

BP_f – ветвящаяся программа;

MBP_f – матричная ветвящаяся программа;

$\widetilde{MBP}_f$ – рандомизированная матричная ветвящаяся программа;

p – большое простое число длиной $\Omega(n) > 512$ бит;

Z_p – кольцо вычетов по модулю p;

L – длина ветвящейся программы BP_f;

W – ширина ветвящейся программы BP_f.

Формально, обфускатор неразличимости iO можно считать компилятором, принимающим на вход булеву функцию $f(x)$ и генерирующим на выходе обфусцированную программу $f'(x) = iO(f(x))$. При этом должно выполняться следующее условие: $\forall x: f'(x) = f(x)$.

Введем понятие неразличимой обфускации для булевых схем класса NC^1, который представляет собой класс булевых функций, вычисляемых схемами из функциональных элементов полиномиальной сложности и логарифмической глубины $O(\log^1 n)$, где n – длина входного вектора.

Определение 1. Вычислительная неразличимость [7].

Пусть $\{X_n\}_n$ и $\{Y_n\}_n$ множества распределений вероятностей над множеством $\{0,1\}^{p(n)}$ для некоторого полинома $p(\cdot)$. Тогда $\{X_n\}_n$ и $\{Y_n\}_n$ вычислительно неразличимы ($\{X_n\}_n \approx \{Y_n\}_n$), если для любой неоднородной полиномиальной вероятностной машины Тьюринга D, существует пренебрежимо малая функция α такая, что $\forall n \in \mathbb{N}$:

$$\Pr[t \leftarrow X_n, D(t) = 1] - \Pr[t \leftarrow Y_n, D(t) = 1] \leq \alpha(n). \quad (1)$$

Определение 2. Обфускатором неразличимости iO называется однородная полиномиальная вероятностная машина Тьюринга для булевых схем класса $\{C_\lambda\}$, где λ – параметр безопасности, если обеспечивается выполнение следующих свойств [4]:

– функциональная эквивалентность. Существует пренебрежимо малая функция $\mu(\lambda)$ такая, что для всех $\lambda \in \mathbb{N}$ и схемы $C \in C_\lambda$:

$$\forall x: \Pr[iO(C(x)) = C(x)] \geq 1 - \mu(\lambda); \quad (2)$$

– свойство виртуального черного ящика (стойкость). Для любой неоднородной полиномиальной вероятностной машины Тьюринга D (распознавателя), существует пренебрежимо малая функция α такая, что для всех $\lambda \in \mathbb{N}$ и любой пары эквивалентных схем $C_1, C_2 \in C_\lambda$, имеющих одинаковый размер, распределения вероятностей случайных величин $iO(C_1)$ и $iO(C_2)$ вычислительно неразличимы, то есть выполняется соотношение:

$$\forall x : \left| \Pr[D(iO(\lambda, C_1(x))) = 1] - \Pr[D(iO(\lambda, C_2(x))) = 1] \right| \le \alpha(\lambda) . \quad (3)$$

Это означает, что не существует алгоритма распознавания обфускации более эффективного, чем обычное предположение, сделанное на основе анализа входов и выходов обфусцированной программы (функции или обфусцированного набора инструкций).

Неразличимость обфусцированных программ означает, что если две эквивалентные программы (одинакового размера) P_1, P_2 ,такие, что $\forall x : P_1(x) = P_2(x)$, а $iO -$ это обфускатор неразличимости, который принимает на вход программу $P(x)$ и на выходе генерирует новую программу $iO(P(x))$, то $iO(P_1(x))$ и $iO(P_2(x))$ будут вычислительно неразличимы:

$$\exists P_1(x), P_2(x). \forall x : P_1(x) = P_2(x) \to iO(P_1(x)) \approx iO(P_2(x)) . \quad (4)$$

Согласно работам [4, 8, 9] процедура неразличимой обфускации включает в себя пять основных этапов.

1. Преобразование булевой функций в вид ветвящейся программы.
2. Преобразование ветвящейся программы в вид матричной ветвящейся программы.
3. Рандомизация матричной ветвящейся программы.
4. Кодирование рандомизированной матричной ветвящейся программы с помощью схемы дифференциального кодирования (Graded Encoding Scheme).
5. Вычисление обфусцированной функции.

Для более подробного рассмотрения аналитической модели неразличимой обфускации на рисунке 1 приведены основные этапы реализации неразличимой обфускации.

$$f \xrightarrow{1} BP_f \xrightarrow{2} MBP_f \xrightarrow{3} M\tilde{B}P_f \xrightarrow{4} iO(f) \xrightarrow{5} f(x)$$

Рис. 1. Основные этапы реализации неразличимой обфускации

Этап 1. Преобразование булевой функции в вид ветвящейся программы. На первом этапе производится преобразование булевой функции $f(x)$ в вид ветвящейся программы BP_f по теореме Сауэр-хоффа [9]. Ветвящиеся программы – это логические схемы, которые хорошо моделируют вычисления с помощью одного процессора, читающего не более одного бита информации в единицу времени.

Определение 3. Ветвящаяся программа для вычисления булевой функции $f(x)$ определяется как ориентированный ациклический граф, вершины которого разделены на слои $0, 1, ..., L$ (L – длина вет-

вящейся программы) таким образом, что выполняются следующие условия:

– из вершин слоя i ребра могут вести только в вершины слоя $i+1$;

– в слое номер 0 имеется единственная вершина, ее входная степень равна нулю;

– каждая вершина слоя L помечена 0 или 1, выходные степени вершин слоя L равны нулю;

– каждая вершина слоев $0,1,...,L-1$ помечена одной из переменных $\{x_1,...,x_n\}$, и имеет два выходящих ребра с пометками 0 и 1.

Процедура вычисления функции $f(x_1,...,x_n)$ с помощью такой программы состоит в том, что осуществляется проход от единственной вершины слоя 0 до некоторой вершины слоя L по ориентированным ребрам графа. Проходя через вершину, помеченную переменной x_i, считывается значение элемента входного вектора $x_i=0$ (или $x_i=1$) и далее осуществляется движение по выходящему ребру с соответствующей пометкой. Добравшись до вершины уровня L, рассчитывается значение функции $f(x_1,...,x_n)$ в соответствии с пометкой на финальной вершине.

Шириной W ветвящейся программы называется максимальное число вершин, находящихся на одном уровне. Размером V ветвящейся программы называется общее количество вершин графа на всех уровнях. На рисунке 2 представлен пример построения ветвящейся программы для вычисления функции $f(x_1,x_2,x_3)=(\overline{x}_1\wedge x_2\wedge\overline{x}_3)\vee(x_1\wedge\overline{x}_2\wedge x_3)$.

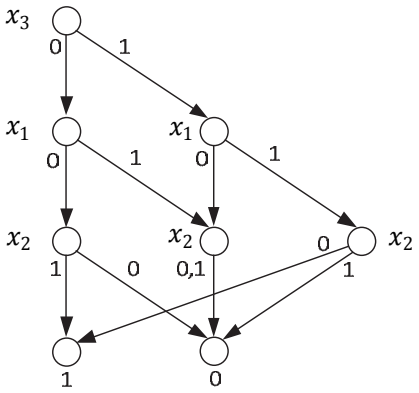


Рис. 2. Ветвящаяся программа для вычисления функции $f(x_1,x_2,x_3)=(\overline{x}_1\wedge x_2\wedge\overline{x}_3)\vee(x_1\wedge\overline{x}_2\wedge x_3)$

В работах [4, 11, 12] для преобразования булевой функции в вид ветвящейся программы применялась теорема Баррингтона. Теорема Баррингтона [13] гласит, что если для булевой функции $f(x)$ существует булева схема глубины L_C, то $f(x)$ можно вычислить ветвящейся программой шириной $W = 5$ и длиной $L \le 4^{L_C}$.

В предлагаемой аналитической модели для преобразования булевой функций в ветвящуюся программу предложено применять теорему Сауэрхоффа. Данный подход позволяет преобразовать любую схему размера V в ветвящуюся программу шириной $W \le 2(V+1)$ и длиной $L \le V$. Теорема Сауэрхоффа более эффективна при преобразовании булевых функций по сравнению с подходом на основе теоремы Баррингтона. Размер ветвящейся программы, полученной по теореме Баррингтона удовлетворяет следующему требованию $V \le L(f)^2$, где $L(f)$ – размер формулы булевой функции, а ветвящаяся программа, построенная по теореме Сауэрхоффа удовлетворяет следующему требованию $V \le 1.360L(f)^{\beta}$, где $\beta = \log_4(3+\sqrt{5}) < 1.195$ [10]. Применение теоремы Сауэрхоффа позволяет значительно сократить объем получаемых обфусцированных данных на выходе аналитической модели.

Этап 2. Преобразование ветвящейся программы в вид матричной ветвящейся программы. На данном этапе ветвящаяся программа BP_f преобразуется в функционально эквивалентную матричную ветвящуюся программу MBP_f.

Определение 4. Пусть $P_{rej} \in \{0,1\}^{W \times W}$ – класс матриц перестановок размерности $W \times W$, таких что $P_{rej} \ne I_{W \times W}$, где $I_{W \times W}$ – единичная матрица размерности $W \times W$. Матричная ветвящаяся программа шириной W и длиной L для входного вектора размером n бит определяется следующей последовательностью:

$$MBP_f = (I_{W \times W}, P_{rej}, inp(i), B_{i,0}, B_{i,1})_{i=1}^{L}, \quad (5)$$

где $B_{i,b} \in \{0,1\}^{W \times W}$ – матрицы перестановок, $b \in \{0,1\}$; $inp(i) : [L] \to [n]$ – функция выбора позиции матриц для текущего входного бита.

Преобразование ветвящейся программы в вид матричной ветвящейся программы осуществляется следующим образом: для каждого слоя $i \in [L]$ ветвящейся программы BP_f составляются перестановки и соответствующие матрицы перестановок $B_{i,0}, B_{i,1}$. При этом выбор матрицы $B_{i,b}$ определяется пометкой b выходящего ребра.

Результат выполнения булевой функции, представленной в виде матричной ветвящейся программы MBP_f, получаемый при входном векторе $x \in \{0,1\}^n$, определяется в соответствии с выражением (6).

$$MBP_f(x) = \begin{cases} 1, \text{ если } \prod_{i=1}^{L} B_{i,inp(i)} = I_{W \times W}; \\ 0, \text{ если } \prod_{i=1}^{L} B_{i,inp(i)} = P_{rej}. \end{cases} \quad (6)$$

Процедура преобразования ветвящейся программы в матричную ветвящуюся программу будет рассмотрена далее на примере перестановочной ветвящейся программы PBP_f фиксированной ширины [14]. На рисунке 3 представлена перестановочная ветвящаяся программа, для которой при входном векторе $x = (0,1,1)$ на выходе вычисляется соответствующая перестановка $P(x) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

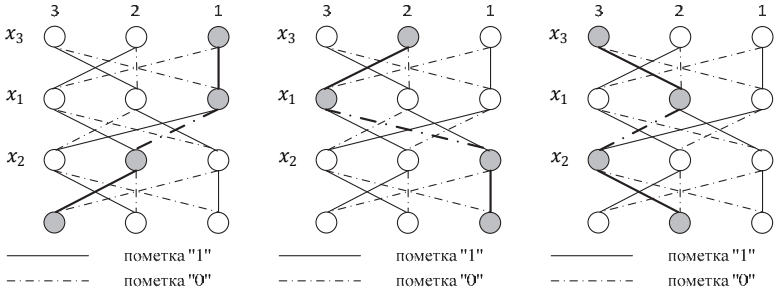


Рис. 3. Перестановочная ветвящаяся программа PBP_f шириной $W = 3$

Для построения матричной ветвящейся программы необходимо построить для первого слоя две перестановки $P_{1,1}, P_{1,0}$:

$$P_{1,1}(x_3) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}; \qquad P_{1,0}(x_3) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

При этом матрицы перестановок $B_{1,1}, B_{1,0}$, соответствующие перестановкам $P_{1,1}, P_{1,0}$ будут иметь следующий вид:

$$B_{1,1=}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix};\qquad B_{1,0=}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}.$$

Затем рассчитываются перестановки и соответствующие им матрицы перестановок для второго и третьего слоя:

$$P_{2,1}(x_1)=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix};\qquad P_{2,0}(x_1)=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix};$$

$$B_{2,1=}\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix};\qquad B_{2,0=}\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix};$$

$$P_{3,1}(x_2)=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix};\qquad P_{3,0}(x_2)=\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix};$$

$$B_{3,1=}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix};\qquad B_{3,0=}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}.$$

Для проверки корректности построенных матриц необходимо рассчитать $\prod_{i=1}^{L} B_{i,inp(i)}$ для входного вектора $x=(0,1,1)$:

$$\prod_{i=1}^{3} B(0,1,1)=B_{1,1}B_{2,0}B_{3,1}=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}. \quad (7)$$

Из выражения (6) следует, что полученная в выражении (7) матрица соответствует исходной перестановке $P(x)=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$. Полученная матрица $\prod_{i=1}^{3} B(0,1,1)=P_{rej}$, соответственно $MBP_f(0,1,1)=0$.

При значении входного вектора $x=(0,1,0)$: $MBP_f(0,1,0)=1$ поскольку:

$$\prod_{i=1}^{3} B(0,1,0) = B_{1,0}B_{2,0}B_{3,1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Этап 3. Рандомизация матричной ветвящейся программы. Для сокрытия базового алгоритма вычисления и значений матриц применяется процедура рандомизации по методу Килиана [15]. Интерпретация протокола скрытой передачи Килиана в аспекте неразличимой обфускации позволила повысить стойкость неразличимой обфускации, в особенности при защите от таких атак как: "атаки с частичным раскрытием", "атаки смешенных входных данных" [4].

Определение 5. Процедура рандомизации матричной ветвящейся программы $rand_p(MBP_f)$ над кольцом Z_p осуществляется следующим образом:

– выбор случайных независимых скалярных значений $\{\alpha_{i,0}, \alpha'_{i,1}, \alpha'_{i,0}, \alpha'_{i,1} \in Z_p : i \in [L]\}$ над кольцом Z_p, таким образом, что $\prod_{i\in I_j} \alpha_{i,0} = \prod_{i\in I_j} \alpha'_{i,0}$ и $\prod_{i\in I_j} \alpha_{i,1} = \prod_{i\in I_j} \alpha'_{i,1}$, где $I_j := \{i \in [L] : inp(i) = j\}$, $j \in [n]$ – набор слоев ветвящейся программы и соответствующий входной бит i;

– для каждого из слоев ветвящейся программы $i \in [L]$ рассчитываются четыре матрицы $D_{i,0}, D_{i,0}, D'_{i,0}, D'_{i,1}$ размерности $(2L+W)\times(2L+W)$ так, что:

$$D_{i,b} = \begin{bmatrix} d_{i,b} & 0 \\ 0 & \alpha_{i,b}B_{i,b} \end{bmatrix}, \qquad D'_{i,b} = \begin{bmatrix} d'_{i,b} & 0 \\ 0 & \alpha'_{i,b}I_{W\times W} \end{bmatrix},$$

где $d_{i,b}, d'_{i,b}$ – случайная диагональная матрица по модулю p размерности $2L\times 2L$, $b \in \{0,1\}$;

– выбор векторов s,t и s',t' размерности $(2L+W)$ таких, что выполняются следующие условия:

$s = (\bar{0}, \bar{s}_R, \hat{s})$, где $\bar{0}$ – нулевой вектор длины L, $\bar{s}_R$ – случайный вектор длины L, $\hat{s}$ – случайный вектор длины W;

$t = (\bar{t}_R, \bar{0}, \hat{t})^T$, где $\bar{0}$ – нулевой вектор длины L, $\bar{t}_R$ – случайный вектор длины L, $\hat{t}$ – случайный вектор длины W;

$s' = (\overline{0}, \overrightarrow{s_R}, \hat{s}')$, где $\overline{0}$ – нулевой вектор длины L, $\overrightarrow{s_R}$ – случайный вектор длины L, $\hat{s}'$ – случайный вектор длины W;

$t' = (\overrightarrow{t_R}, \overline{0}, \hat{t}')^T$, где $\overline{0}$ – нулевой вектор длины L, $\overrightarrow{t_R}$ – случайный вектор длины L, $\hat{t}$ – случайный вектор длины W;

$\langle \hat{s}, \hat{t} \rangle = \langle \hat{s}', \hat{t}' \rangle$ – скалярные произведения элементов векторов s, t и s', t' равны.

– выбор $2(L+1)$ произвольных невырожденных матриц $R_0, R_1, ..., R_L, R'_0, R'_1, ..., R'_L \in Z_p^{(2L+W)(2L+W)}$;

– вычисление матриц $\widetilde{D}_{i,b} \coloneqq R_{i-1} D_{i,b} (R_i)^{-1}$ и $\widetilde{D}'_{i,b} \coloneqq R'_{i-1} D'_{i,b} (R'_i)^{-1} \quad \forall i \in [n], b \in \{0,1\}$;

– вычисление рандомизированных значений векторов $\widetilde{s} = s \cdot R_0^{-1}$; $\widetilde{t} = R_L \cdot t$ и $\widetilde{s}' = s' \cdot (R'_0)^{-1}$; $\widetilde{t}' = R'_L \cdot t'$.

Таким образом, результатом данного этапа является рандомизированная матричная ветвящаяся программа над кольцом Z_p, которая представляет собой следующую совокупность:

$$\widetilde{MBP}_f(x) = \left\{ \begin{array}{c} \widetilde{s} = s \cdot R_0^{-1}, \ \widetilde{t} = R_L \cdot t \\ \{\widetilde{D}_{i,b} = R_{i-1} D_{i,b} R_i^{-1}\}_{\forall i \in [L], b \in \{0,1\}} \\ \widetilde{s}' = s' \cdot (R'_0)^{-1}, \ \widetilde{t}' = R'_L \cdot t' \\ \{\widetilde{D}'_{i,b} \coloneqq R'_{i-1} D'_{i,b} (R'_i)^{-1}\}_{\forall i \in [L], b \in \{0,1\}} \end{array} \right\}.$$

Этап 4. Кодирование рандомизированной матричной ветвящейся программы с помощью схемы дифференциального кодирования. Для кодирования рандомизированных матриц ветвящейся программы можно применяться схема дифференциального кодирования, которая описана в [16, 17]. Предложенная схема обладает свойством гомоморфного шифра [18]: обеспечивает конфиденциальность значений каждой из матриц при возможности осуществления математических операций над закодированными данными без их раскодирования. Разрешенными операциями над зашифрованными данными при этом являются произведение и сложение, которые удовлетворяют следующему соотношению:

$$\begin{cases} E(m_1)\otimes E(m_2)=E(m_1\cdot .m_2), \\ E(m_1)\oplus E(m_2)=E(m_1+m_2); \end{cases}$$

где $\otimes$ и $\oplus$ – операции над шифрованными данными, соответствующие операциям умножения и сложения над открытыми данными; $E(m_1), E(m_2)$ – шифр тексты сообщений m_1 и m_2 соответственно.

Однако в отличие от полностью гомоморфного шифрования в схеме дифференциального кодирования существуют ограничения на операции сложения и умножения. Каждая зашифрованная матрица кодируется относительно некоторого множества, являющегося подмножеством универсума $S \in [U]$. При этом две матрицы может складывать только, если они кодируются по отношению к одному множеству S. В свою очередь умножение возможно только, если шифрование осуществляется относительно двух непересекающихся множеств S и S'.

Определение 6. Система разделенных множеств представляет собой совокупность множеств $S_n = \{S_{i,b} : i \in [n], b \in \{0,1\}\}$ из $2n$ элементов над универсумом $U = \{1,2,...,U_{\max}\}$, таких что $\bigcup_{i\in[n]} S_{i,0} = \bigcup_{i\in[n]} S_{i,1} = U$, $S_{i,0} = \bigcup_{k=1}^{|ind(i)|} S_{i,0}^k$, $S_{i,1} = \bigcup_{k=1}^{|ind(i)|} S_{i,1}^k$ при этом $\forall i_1, i_2, k_1, k_2 : S_{i_1,b}^{k_1} \cap S_{i_2,b}^{k_2} = \varnothing$, $b \in \{0,1\}$, $|ind(i)|$ – частота встречаемости i-го входного бита в $M\tilde{B}P_f$ программе длиной L. Значение $U_{\max}$ рассчитывается по следующей формуле: $U_{\max} = \sum_{i=1}^{n} (2 \cdot |ind(i)| - 1)$.

Предложенная система разделенных множеств позволяет осуществить полное покрытие универсума U в случае корректного умножения матриц – структура ветвящейся программы хранит взаимосвязи входных бит и слоев программы (соответствующих матриц). В выражении (8) представлен предлагаемый подход к построению системы разделенных множеств в рамках одного i-го входного бита при $k = |ind(i)|$:

$$\begin{aligned} &S_{1,0} = \{1\}, S_{2,0} = \{2,3\}, ..., S_{k-1,0} = \{2k-4, 2k-3\}, S_{k,0} = \{2k-2, 2k-1\}; \\ &S_{1,1} = \{1,2\}, S_{2,1} = \{3,4\}, ..., S_{k-1,1} = \{2k-3, 2k-2\}, S_{k,1} = \{2k-1\}. \end{aligned} \quad (8)$$

Пример построения системы разделенных множеств представлен на рисунке 4.

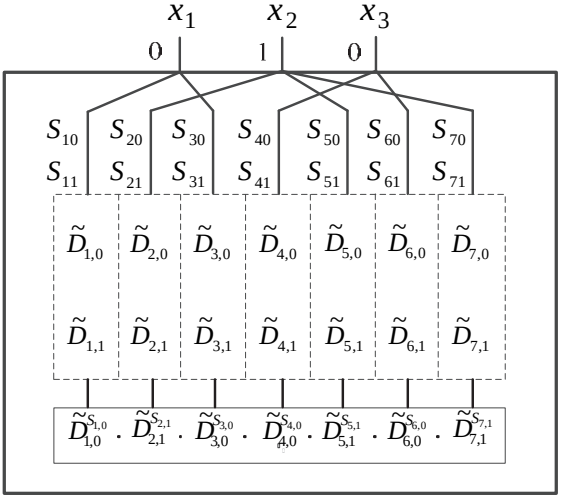


Рис. 4. Пример построения системы разделенных множеств

Пусть на вход ветвящейся программы длиной $L = 7$ поступает вектор $x = (0,1,0)$. При этом $|ind(x_1)| = 2$, $|ind(x_2)| = 3$, $|ind(x_3)| = 2$. Построение системы осуществляется следующим образом: последовательно для каждого входного бита задаются элементы подмножеств S таким образом, чтобы выполнялось условие $\bigcup_{i\in[n]} S_{i,0} = \bigcup_{i\in[n]} S_{i,1} = U$:

$$x_1:\ S_{1,0} = \{1\}, S_{3,0} = \{2,3\}, S_{1,1} = \{1,2\}, S_{3,1} = \{3\};$$

$$x_2:\ S_{2,0} = \{4\}, S_{5,0} = \{5,6\}, S_{7,0} = \{7,8\}, S_{2,1} = \{4,5\}, S_{5,1} = \{6,7\}, S_{7,1} = \{8\};$$

$$x_3:\ S_{4,0} = \{9\}, S_{6,0} = \{10,11\}, S_{4,1} = \{9,10\}, S_{6,1} = \{11\}.$$

Предложенная система разделенных множеств позволяет защититься от атак с использованием смешанных произведений. Злоумышленнику не известны взаимосвязи входных бит и соответствующих матриц, за счет осуществления смешанных произведений он может попытаться восстановить эти взаимосвязи. Но без их точного знания злоумышленник не сможет осуществить полное покрытие универсума U, а тем самым и вычисление зашифрованной функции.

Определение 7. Для $k+1$ мультипликативных циклических групп $G_1,...,G_k,G_T$, конечного порядка p, существует k–линейное отображение $e: G_1 \times ... \times G_k \to G_T$, если выполняются следующие свойства:

– для всех элементов $g_1 \in G_1, ..., g_k \in G_k, i \in [k]$ и целого $\alpha \in Z_p$:

$$e(g_1, ..., \alpha \cdot g_i, ..., g_k) = \alpha \cdot e(g_1, ..., g_k);$$

– если элементы $g_i \in G_i, i \in [k]$, являются образующими элементами групп G_i, то $e(g_1, ..., \alpha \cdot g_i, ..., g_k)$ является образующим элементом группы G_T.

В настоящие время известны два наиболее общих подхода к построению k–линейных отображений, реализующих схему дифференциального кодирования. Первый из них [15] базируется на полилинейных отображениях на решетках, а второй [16] – для целых чисел. Предлагаемая аналитическая модель базируется на втором подходе, поскольку реализация полилинейных отображений на основе целых чисел является более эффективной в сравнении с применением решеток [16].

Определение 8. Схемой дифференциального кодирования называется совокупность вероятностных полиномиальных алгоритмов $(InstGen, Enc, Add, Mult, isZero)$, которые выполняют следующие функции:

– Instance Generation: $(sp, pp) \leftarrow InstGen(1^\lambda, 1^k)$.

$InstGen$ принимает на вход λ (параметр безопасности) и k (порядок полилинейности отображения). На выходе генерируется секретный параметр sp и открытый параметр pp. Секретный параметр sp содержит целое число Y, такое что $k \le Y \le 2k$, простые числа $g_1, ..., g_N$, где $N = \eta \log_2 \lambda$, $\eta = (k+1)(5\lambda+2)+6$ [16] и совокупность множеств $\{E_S^m : m \in Z_{g_1} \times ... \times Z_{g_N}, S \subseteq [U]\}$. E_S^m рассматривается как набор возможных кодирований значений m относительно множества S.

– Encoding: $u \leftarrow Enc(sp, m, S)$.

Enc принимает на входе секретный параметр sp, значение открытого сообщения $m \in Z_{g_1} \times ... \times Z_{g_N}$, и множество $S \subseteq [U]$, на выходе

осуществляется шифрование m относительно множества S (обозначение – $u \in E_S^m$).

– Addition: $u \leftarrow Add(pp, u, u')$.

Add принимает на входе открытый параметр pp и зашифрованные сообщения $u \in E_S^m$, $u' \in E_{S'}^{m'}$, на выходе вычисляется значение зашифрованного сообщения равного сумме двух исходных зашифрованных сообщений $u \in E_S^{m+m'}$, в том случае, если $S = S'$.

– Multiplication: $u \leftarrow Mult(pp, u, u')$.

$Mult$ принимает на входе открытый параметр pp и зашифрованные сообщения $u \in E_S^m$, $u' \in E_{S'}^{m'}$, на выходе вычисляется значение зашифрованного сообщения равного произведению двух исходных зашифрованных сообщений $u \in E_{S \bigcup S'}^{m \cdot m'}$, в том случае, если $S \bigcap S' = \varnothing$.

– Zero Test: $b \leftarrow isZero(pp, u)$.

$isZero$ принимает на входе открытый параметр pp и зашифрованное сообщение u, на выходе вычисляется значение, характеризующее открытое сообщение, 1 – в случае, если $u \in E_{[U]}^0$ (кодирование нуля), и 0 – в остальных случаях.

Далее представлено описание схемы дифференциального кодирования для шифрования рандомизированной матричной ветвящейся программы в рамках аналитической модели. Для программы длиной L необходима реализация $(L+2)$ – линейного отображения, поскольку L – число умножений элементов входного вектора в программе, но также необходимо учесть умножение на вектора s, t. При этом основные операции схемы дифференциального кодирования принимают следующий вид:

– $InstGen$: пусть $Y = L + 2$ – требуемый размер системы разделенных множеств. Изначально выбираются N секретных больших простых чисел $\{p_i\}_{i=1}^N$, затем вычисляется их произведение $x_0 = \prod_{i=1}^N p_i$. Также выбираются N секретных простых чисел $\{g_i\}_{i=1}^N$, N случайных целых чисел $\{h_i\}_{i=1}^N$, N случайных целых чисел $\{r_i\}_{i=1}^N$ и Y случайных значений $z_1, \ldots, z_Y \in Z_{x_0}$. Затем осуществляется вы-

числение параметра процедуры Zero Test p_{zt}, заданного целым числом, в соответствии с выражением (9):

$$p_{zt} = \sum_{i=1}^{N} h_i.(\prod_{j=1}^{z} z_j.g_i^{-1} \bmod p_i).\prod_{i' \neq i} p_{i'} \bmod x_0 \ . \qquad (9)$$

Выходными для данной процедуры являются следующие параметры: секретные – $(\{z_i\}_{i=1}^{Y}, \{g_i\}_{i=1}^{N}, \{p_i\}_{i=1}^{N})$, открытые – (p_{zt}, x_0).

– *Enc*: пусть $u \in Z_{x_0}$ является кодированием сообщения $m = (m_1,...,m_N) \in Z_{g_1} \times ... \times Z_{g_N}$ относительно множества $S \subseteq [U]$, тогда для всех $1 \le i \le N$ и случайных (малых) целых чисел r_i:

$$u \equiv \frac{r_i.g_i + m_i}{\prod_{j \in S} z_j} (\bmod p_i) .$$

– *Add*: при заданных $u, u' \in Z_{x_0}$, где

$$\forall i : u \equiv \frac{r_i.g_i + m_i}{\prod_{j \in S} z_j} (\bmod p_i), \qquad u' \equiv \frac{r_i'.g_i + m_i'}{\prod_{j \in S} z_j} (\bmod p_i) ,$$

справедливо следующее выражение:

$$\forall i : u + u' \equiv \frac{(r_i + r_i').g_i + (m_i + m_i')}{\prod_{j \in S} z_j} (\bmod p_i) .$$

– *Mult*: Пусть S и S' являются множествами такими, что $S \bigcap S' = \varnothing$. Тогда при заданных $u, u' \in Z_{x_0}$, где

$$\forall i : u \equiv \frac{r_i.g_i + m_i}{\prod_{j \in S} z_j} (\bmod p_i), \qquad u' \equiv \frac{r_i'.g_i + m_i'}{\prod_{j \in S'} z_j} (\bmod p_i) ,$$

справедливо следующее выражение:

$$\forall i : u \cdot u' \equiv \frac{(r_i r_i' + r_i m_i' + r_i' m_i) g_i + m_i m_i'}{\prod_{j \in S \bigcup S'} z_j} (\bmod p_i) ,$$

где $\forall i : (r_i r_i' + r_i m_i' + r_i' m_i) g_i + m_i m_i' \le p_i$.

Согласно представленной схеме для дифференциального кодирования рандомизированной матрицы ветвящейся программы $\widetilde{MBP}_f$ необходима реализация $(L+2)$ – уровневой схемы кодирования. При этом также необходимо сгенерировать:

– r_s, r_t, r'_s, r'_t – случайные вектора, которые используются для кодирования векторов $\tilde{s}, \tilde{t}, \hat{s}', \hat{t}'$;

– $U_{i,b}, U'_{i,b}$ – случайные матрицы ($i \in [L], b \in \{0,1\}$).

Пусть $U_s, U_t, U_1, ..., U_L$ – непересекающиеся множества универсума U такие, что $U = U_s \cup U_t \cup \bigcup_{j=1}^{L} U_j$. Аналогично и для множеств, образующих универсум $U = U'_s \cup U'_t \cup \bigcup_{j=1}^{n} U'_j$. При этом U_s, U_t и U'_s, U'_t используются для кодирования $\tilde{s}, \tilde{t}$ и $\tilde{s}', \tilde{t}'$ соответственно.

Таким образом, результатом данного этапа является обфусцированная функция $iO(f(x))$, которая представляет собой следующую совокупность:

$$iO(f(x)) = \left\{ \begin{array}{c} s^* = [(z_0^{-1}(\tilde{s} + gr_s))_{x_0}]_{U_s} \\ (s')^* = [(z_0^{-1}(\tilde{s}' + gr'_s))_{x_0}]_{U'_s} \\ t^* = [(z_{L+1}^{-1}(\tilde{t}' + gr_t))_{x_0}]_{U_t} \\ (t')^* = [(z_{L+1}^{-1}(\tilde{t}' + gr'_t))_{x_0}]_{U'_t} \\ \{[D^*_{i,b} = [z_i^{-1}(\tilde{D}_{i,b} + g \cdot U_{i,b})]_{x_0}]_{S(i,b)}\}_{i \in [L], b \in \{0,1\}} \\ \{[(D'_{i,b})^* = [z_i^{-1}(\tilde{D}'_{i,b} + g \cdot U'_{i,b})]_{x_0}]_{S'(i,b)}\}_{i \in [L], b \in \{0,1\}} \end{array} \right\},$$

где $g = CRT(g_i), g_i \le p_i$ – вычисляется по китайской теореме об остатках [19].

Этап 5. Вычисление обфусцированной булевой функции.

Вычисление значения обфусцированной булевой функций $iO(f(x))$ при заданном входном векторе x осуществляется с помощью процедур *Add* и *Mult*, которые в свою очередь позволяют вычислить значение выражения (10). Данное выражение представляет собой кодирование разности произведений рандомизированных матриц $\tilde{D}$ и $\tilde{D}'$:

$$Enc(\tilde{s}.\prod_{i=1}^{L} \tilde{D}_{i,inp(i)}.\tilde{t} - \tilde{s}'.\prod_{i=1}^{L} \tilde{D}'_{i,inp(i)}.\tilde{t}') = u. \qquad (10)$$

Пусть

$$q = \tilde{s} \cdot \prod_{i=1}^{L} \tilde{D}_{i,inp(i)} \cdot \tilde{t} = \tilde{s} \cdot (R_0 D R_L^{-1}) \cdot \tilde{t}^T = \hat{s} \cdot B \cdot \hat{t}^T,$$

$$q' = \tilde{s}' \cdot \prod_{i=1}^{n} \tilde{D}'_{i,inp(i)} \cdot \tilde{t}' = \tilde{s}' \cdot (R_0 D' R_L^{-1}) \cdot \tilde{t}'^T = \hat{s}' \cdot I_{W \times W} \cdot \hat{t}'^T.$$

Если B является единичной матрицей, то согласно выражению (6) и условию $\langle \hat{s}, \hat{t} \rangle = \langle \hat{s}', \hat{t}' \rangle$:

$$q - q' = \hat{s} \cdot B \cdot \hat{t}^T - \hat{s}' \cdot I_{W \times W} \cdot \hat{t}'^T = \hat{s} \cdot I_{W \times W} \cdot \hat{t}^T - \hat{s}' \cdot I_{W \times W} \cdot \hat{t}'^T = 0. \qquad (11)$$

Из выражения (11) следует, что $u = Enc(q - q') = Enc(0)$ является кодированием 0.

Для непосредственного вычисления значения обфусцированной булевой функции применяется операция $isZero$. Если $isZero(pp, u) = 1$, то $f(x) = 1$, иначе $f(x) = 0$. На рисунке 5 представлена иллюстрация процедуры вычисления обфусцированной функции при определенном входном векторе x.

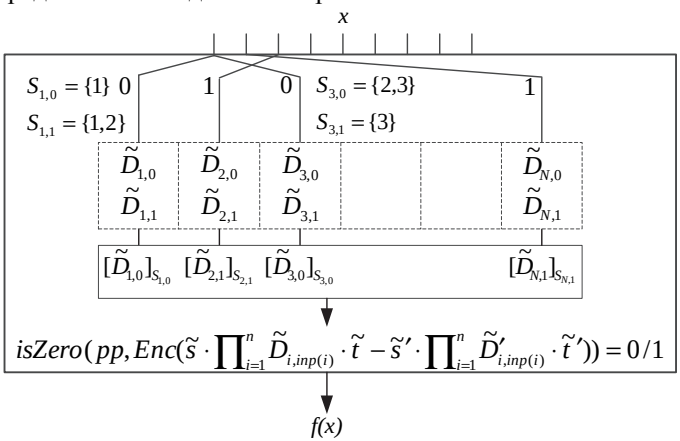


Рис 5. Процедура вычисления обфусцированной булевой функции при определенном входном векторе x

Если u является кодированием нуля $\forall i \in [N]: m_i = 0$, то, исходя из предыдущего выражения, следует:

$$p_{zt} \cdot u = \sum_{i=1}^{N} h_i r_i \cdot \prod_{i' \neq i} p_{i'} (\mathrm{mod}\, x_0). \qquad (13)$$

Поскольку $h_i.r_i << p_i$, то $h_i.r_i.p_{i'} = h_i.r_i \cdot \frac{x_o}{p_i} << x_0$, а также сумма $\sum_{i=1}^{N} h_i r_i p_{i'} << x_0$. Из выражения (12) следует, что $\omega = p_{zt} \cdot u (\mathrm{mod}\ x_0)$ имеет малое значение по сравнению с x_0, тогда произведения h_i, r_i должны удовлетворять определенным ограничениям по размеру. Поэтому согласно лемме 3 из [16], при $m = 0$ – $|\omega| < x_0 \cdot 2^{-\upsilon-\lambda-2}$, а если $m \neq 0$, то $|\omega| \geq x_0 \cdot 2^{-\nu+2}$, где $\upsilon = \alpha + \beta + 11$, α – размерность в битах простых чисел $\{g_i\}_{i=1}^{N}$, β – размерность в битах случайных целых чисел $\{h_i\}_{i=1}^{N}$. Таким образом:

$$isZero(\ pp,u\ ) = \begin{cases} 1,\ если\ |\omega| < x_0.2^{-\upsilon-\lambda-2}; \\ 0,\ если\ |\omega| \geq x_0.2^{-\upsilon+2}. \end{cases} \quad (14)$$

Из выражения (14) следует, что значение обфусцированной функции будет вычисляться следующим образом:

$$O(\ f(\ x\ )) = \begin{cases} 1, если\ isZero = 1; \\ 0, если\ isZero = 0. \end{cases} \quad (15)$$

4. Оценка свойств аналитической модели защиты файлов документальных форматов от несанкционированного доступа. Разработанная модель предназначена для описания этапов, необходимых для осуществления неразличимой обфускации программного кода, с целью защиты файлов документальных форматов от несанкционированного доступа.

Согласно общепринятым требованиям модель должна удовлетворять следующим свойствам: адекватность (соответствие модели исходной реальной системе и учет, наиболее важных качеств, связей и характеристик), точность (степень совпадения полученных в процессе моделирования результатов с характеристиками реального объекта), универсальность (применимость модели к анализу ряда однотипных систем в одном или нескольких режимах функционирования), целесообразность (точность получаемых результатов и общность решения задачи должны увязываться с затратами на моделирование) [20].

Адекватность. Разработанная модель позволяет адекватно представить все элементы, необходимые для осуществления неразли-

чимой обфускации программного кода, определить требования к ним так, чтобы они удовлетворяли требованиям по стойкости и функциональной эквивалентности обфусцированной программы.

Точность. Точность представленной модели характеризуется выполнением предъявленных к ней требований по стойкости и функциональной эквивалентности обфусцированной программы.

Универсальность. Представленная модель подходит для всех видов файлов документальных форматов и любых возможных видов угроз.

Целесообразность. Для осуществления защиты файлов документальных форматов от несанкционированного доступа на основе математического аппарата неразличимой обфускации представленная модель является необходимой.

Экспериментальная проверка модели и формальное доказательство стойкости разработанного подхода является направлением дальнейших исследований.

5. Заключение. Разработанная аналитическая модель защиты от несанкционированного доступа к файлам документальных форматов, отличающаяся применением процедуры неразличимой обфускации программного кода позволяет реализовать метод контролируемого разграничения доступа в автоматизированной системе, обрабатывающей информацию различного уровня конфиденциальности, при условии разрешения одновременного доступа пользователя к ресурсам различного уровня конфиденциальности.

Литература

1. *Козачок А.В., Туан Л.М.* Обоснование возможности применения неразличимой обфускации для защиты исполняемых файлов // Перспективные информационные технологии: Сб. тр. междунар. НТК. Самара: 2015. Т. 1. С. 269–272.
2. *Козачок А.В., Туан Л.М.* Комплекс алгоритмов контролируемого разграничения доступа к данным, обеспечивающий защиту от несанкционированного доступа // Системы управления и информационные технологии. 2015. № 3(61). С. 58–61.
3. *Варновский Н.П., Захаров В.А., Кузюрин Н.Н., Шокуров А.А.* Современное состояние исследований в области обфускации программ: определения стойкости обфускации // Труды Института Системного программирования: М.: ИСП РАН. 2014. Т. 26. № 3. С. 167–198.
4. *Garg S., Gentry C., Halevi S., Raykova M., Sahai A., Waters B.* Candidate indistinguishability obfuscation and functional encryption for all circuits // In FOCS. 2013. pp. 40–49.
5. *Варновский Н.П., Захаров В.А., Кузюрин Н.Н.* Математические проблемы обфускации // Математика и безопасность информационных технологий. Материалы конференции в МГУ. Москва: 2005. С. 65–91.
6. *Barak B., Goldreich O., Impagliazzo R., Rudich S., Sahai A., Vadhan S.P., Yang K.* On the (im)possibility of program obfuscation // In Advances in Cryptology – CRYPTO. 2001. pp. 1–18.

7. *Pass R., Shelat A.* A Course in Cryptograph // Theoretical Foundation of Cryptography. 2010. pp. 68–71.
8. *Ananth P., Gupta D., Ishai Y., Sahai A.* Optimizing obfuscation: Avoiding barrington's theorem // In Proceeding of the 2014 ACM SIGSAC. 2014. pp. 646–658.
9. *Pass R., Seth K., Telang S.* Indistinguishability obfuscation from semantically-secure multilinear encodings // In Advances in Cryptology – CRYPTO. 2014. pp. 500–517.
10. *Sauerhoff M., Wegener I., Werchner R.* Relating branching program size and formula size over the full binary basis // In Proceedings of 16th STACS. 1999. pp. 57–67.
11. *Barak B., Garg S., Kalai Y. T., Paneth O., Sahai A.* Protecting obfuscation against algebraic attacks // In Advances in Cryptology – EUROCRYPT. 2014. pp. 221–238.
12. *Brakerski Z., Rothblum G. N.* Virtual black-box obfuscation for all circuits via generic graded encoding // In Proceedings of 11th Theory of Cryptography Conference, TCC. 2014. pp. 1–25.
13. *Barrington D. A.* Bounded-width polynomial-size branching programs recognize exactly those languages in NC1 // Journal of Computer and System Sciences. 1989. pp. 150–164.
14. *Balaji N., Krebs A., Limaye N.* Skew Circuits of Small Width // In Proceedings of 21st International Conference, COCOON. Beijing. 2015. pp. 199–210.
15. *Kilian J.* Founding cryptography on oblivious transfer // In 20th Annual ACM Symposium on Theory of Computing. 1988. pp. 20–31.
16. *Garg S., Gentry C., Halevi S.* Candidate multilinear maps from ideal lattices // In Advances in Cryptology – EUROCRYPT. 2013. pp. 1–17.
17. *Coron J. S., Lepoint T., Tibouchi M.* Practical multilinear maps over the integers // In In Advances in Cryptology – CRYPTO 2013. pp. 476–493.
18. *Варновский Н.П., Шокуров А.В.* Гомоморфное шифрование // Труды Института Системного программирования: М.: ИСП РАН. 2006. Т. 12. С. 27–36.
19. *Нестеренко Ю.В.* Теория чисел // М.: Академия. 2008. 273 с.
20. *Бабенко Л.К., Буртыка Ф.Б., Макаревич О.Б., Трепачева А.В.* Обобщенная модель системы криптографически защищенных вычислений // Известия ЮФУ. Технические науки. 2015. № 5(166). С. 77–86.

References

1. Kozachok A.V., Tuan L.M. [On the possibility of application indistinguishability obfuscation to protect executable files] *Perspektivnye informacionnye tehnologii: Sb. tr. mezhdunar. NTK* [Advanced information technologies: Proceedings of the International Conference]. Samara: 2015. vol. 1. pp. 269–272. (In Russ.).
2. Kozachok A.V., Tuan L.M. [Controlling data access restriction algorithm providing protection from unauthorized access] *Sistemy upravlenija i informacionnye tehnologii – Control Systems and Information Technologies*. 2015. vol. 3(61). pp. 58–61. (In Russ.).
3. Varnovskij N.P., Zaharov V.A., Kuzjurin N.N., Shokurov A.V. [Modern state of research in program obfuscation: the definition of obfuscation security]. *Trudy Instituta Sistemnogo programmirovanija – Proceedings of the Institute for System Programming*. M.: ISP RAS. 2014. vol. 26(3). pp. 167–198. (In Russ.).
4. Garg S., Gentry C., Halevi S., Raykova M., Sahai A., Waters B. Candidate indistinguishability obfuscation and functional encryption for all circuits. *In FOCS*. 2013. pp. 40–49.
5. Varnovskij N.P., Zaharov V.A., Kuzjurin N.N. [Mathematical problems of obfuscation]. *Matematika i bezopasnost' informacionnyh tehnologij. Materialy konferencii v MGU* [Mathematics and security of information technology. Proceedings of the conference at Moscow State University]. Moscow: 2005. pp. 65–91. (In Russ.).

6. Barak B., Goldreich O., Impagliazzo R., Rudich S., Sahai A., Vadhan S.P., Yang K. On the (im)possible obfuscating programs. In Advances in Cryptology – CRYPTO. 2001. pp. 1–18.
7. Pass R., Shelat A. A Course in Cryptograph. Theoretical Foundation of Cryptography. 2010. pp. 68–71.
8. Ananth P., Gupta D., Ishai Y., Sahai A. Optimizing obfuscation: Avoiding barrington's theorem. In Proceeding of the 2014 ACM SIGSAC. 2014. pp. 646–658.
9. Pass R., Seth K., Telang S. Indistinguishability obfuscation from semantically-secure multilinear encodings. In Advances in Cryptology – CRYPTO. 2014. pp. 500–517.
10. Sauerhoff M., Wegener I., Werchner R. Relating branching program size and formula size over the full binary basis. In Proceedings of 16th STACS. 1999. pp. 57–67.
11. Barak B., Garg S., Kalai Y. T., Paneth O., Sahai A. Protecting obfuscation against algebraic attacks. In Advances in Cryptology – EUROCRYPT 2014. pp. 221–238.
12. Brakerski Z., Rothblum G. N. Virtual black-box obfuscation for all circuits via generic graded encoding. In Proceedings of 11th Theory of Cryptography Conference, TCC. 2014. pp. 1–25.
13. Barrington D. A. Bounded-width polynomial-size branching programs recognize exactly those languages in NC1. *Journal of Computer and System Sciences*. 1989. pp. 150–164.
14. Balaji N., Krebs A., Limaye N. Skew Circuits of Small Width. In Proceedings of 21st International Conference, COCOON. Beijing. 2015. pp. 199–210.
15. Kilian J. Founding cryptography on oblivious transfer. In 20th Annual ACM Symposium on Theory of Computing. 1988. pp. 20–31.
16. Garg S., Gentry C., Halevi S. Candidate multilinear maps from ideal lattices. In Advances in Cryptology – EUROCRYPT. 2013. pp. 1–17.
17. Coron J. S., Lepoint T., Tibouchi M. Practical multilinear maps over the integers. In In Advances in Cryptology – CRYPTO. 2013. pp. 476–493.
18. Varnovskij N.P., Shokurov A.V. [Homomorphic encryption]. *Trudy Instituta Sistemnogo programmirovanija – Proceedings of the Institute for System Programming*. M.: ISP RAS. 2006. vol. 12. pp. 27–36. (In Russ.).
19. Nesterenko Ju.V. *Teorija chisel* [Number theory]. M. Academy. 2008. 273 p. (In Russ.).
20. Babenko L.K., Burtyka F.B., Makarevich O.B., Trepacheva A.V. [The generalized model of a cryptographically secure computing]. *Izvestiya YuFU. Tekhnicheskie nauki – Izvestiya SFedU. Engineering Sciences*. 2015. vol. 5(166), pp. 77-86. (In Russ.).

Козачок Александр Васильевич — к-т техн. наук, сотрудник, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: информационная безопасность, защита от несанкционированного доступа, математическая криптография, теоретические проблемы информатики. Число научных публикаций — 68. alex.totrin@gmail.com; Приборостроительная, 35, Орел, 302034; р.т.: +7(486) 254-99-33.

Kozachok Alexander Vasilevich — Ph.D., researcher, The Academy of Federal Security Guard Service of the Russian Federation. Research interests: information security, unauthorized access protection, mathematical cryptography, theoretical problems of computer science. The number of publications — 68. alex.totrin@gmail.com; 35, Priborostroitelnaya Street, Orel, 302034, Russia; office phone: +7(486) 254-99-33.

Бочков Максим Вадимович — д-р техн. наук, заместитель директора по научной и учебной работе, частное образовательное учреждение дополнительного профессионального образования «Центр предпринимательских рисков» (ЧОУ ДПО "ЦПР"). Область научных интересов: информационная безопасность, защита от несанкционированного

доступа, защита информации в автоматизированных системах. Число научных публикаций — 150. mvboch@yandex.ru; ул. Профессора Попова, 27, Санкт-Петербург, 197022; р.т.: +7 (812) 234-95-66.

Bochkov Maksim Vadimovich — Ph.D., Dr. Sci., professor, deputy director for science, Business risk educational center. Research interests: information security, unauthorized access protection. The number of publications — 150. mvboch@yandex.ru; 27, Professor Popov Street, Saint-Petersburg, 197022, Russia; office phone: +7 (812) 234-95-66.

Фаткиева Роза Равильевна — к-т техн. наук, доцент, старший научный сотрудник лаборатории информационно-вычислительных систем и технологии программирования, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: моделирование информационных систем. Число научных публикаций — 50. rrf@iias.spb.su; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; р.т.: +7-812-328-43-69, Факс: +7 (812)350-1113.

Fatkieva Roza Ravilievna — Ph.D., associate professor, senior researcher of computer and information systems and software engineering laboratory, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: modeling of information systems. The number of publications — 50. rrf@iias.spb.su; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7-812-328-43-69, Fax: +7 (812)350-1113.

Туан Лай Минь — сотрудник, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: информационная безопасность, защита от несанкционированного доступа. Число научных публикаций — 8. lmtuan.1989@gmail.com; Приборостроительная, 35, Орел, 302034; р.т.: +7(486) 254-99-33.

Tuan Lai Minh — researcher, The Academy of Federal Security Guard Service of the Russian Federation. Research interests: information security, unauthorized access protection. The number of publications — 8. lmtuan.1989@gmail.com; 35, Priborostroitelnaya Street, Orel, 302034, Russia; office phone: +7(486) 254-99-33.

РЕФЕРАТ

Козачок А.В., Бочков М.В., Фаткиева Р.Р., Туан Л.М. **Аналитическая модель защиты файлов документальных форматов от несанкционированного доступа.**

Целью проводимого исследования является построение комплекса моделей и алгоритмов процесса контролируемого разграничения доступа к файлам документальных форматов, позволяющего осуществить защиту от несанкционированного доступа к информации за счет применения неразличимой обфускации программного кода.

В статье подробно рассмотрено описание аналитической модели неразличимой обфускации программного кода, положенной в основу предлагаемого подхода к защите от несанкционированного доступа к файлам документальных форматов.

Процедуру проверки прав доступа пользователя к документу, внедренному в контейнер, можно рассматривать как точечную функцию, поскольку результатом ее выполнения является значение из множества {0,1}, поэтому применение неразличимой обфускации для защиты данной проверки является также корректным. Данная работа посвящена модификации существующих подходов к осуществлению неразличимой обфускации с целью устранения ряда ограничений, обусловленных применяемыми механизмами, моделями и алгоритмами.

Разработанная модель неразличимой обфускации включает в себя пять основных этапов.

1. Преобразование булевой функций в вид ветвящейся программы.
2. Преобразование ветвящейся программы в вид матричной ветвящейся программы.
3. Рандомизация матричной ветвящейся программы.
4. Кодирование рандомизированной матричной ветвящейся программы с помощью схемы дифференциального кодирования.
5. Вычисление обфусцированной функции.

Отличительными особенностями разработанной модели являются:

— применение теоремы Сауэрхоффа для преобразования булевой функции в вид ветвящейся программы, что позволило добиться уменьшения длины формируемой ветвящейся программы;

— использование системы разделенных множеств для защиты от атак "смешенных произведений" и "смешанных входных данных";

— уточнение ряд процедур, осуществляемых на различных этапах неразличимой обфускации программного кода.

SUMMARY

Kozachok A.V., Bochkov M.V., Fatkieva R.R., Tuan L.M. **Analytical Model for Protecting Documentary File Formats from Unauthorized Access.**

The purpose of the research is to construct a complex of models and algorithms to control data access restriction to the documentary file formats based on indistinguishable program code obfuscation.

The article discussed in detail the description of the indistinguishable code obfuscation analytical model that underlies the proposed approach to protect against unauthorized access to the files of documentary formats.

The process of checking user access rights to the document, encapsulated in a container, may be considered as a point function because its value is from the set {0, 1}. So using indistinguishable obfuscation to protect this process is correct. This work is devoted to the modification of existing approaches to indistinguishable obfuscation and removing some restrictions, imposed by the applicable mechanisms, models and algorithms.

The developed model of indistinguishable obfuscation consists of five basic steps.

1. Convert Boolean function to branching program.
2. Convert branching program to matrix branching program.
3. Randomization of matrix branching program.
4. Encoding randomized matrix branching program with graded encoding scheme.
5. Evaluation of obfuscated function.

The distinctive features of the model are:

— the application of Sauerhoff Theorem to convert a Boolean function into the form of branching programs, which resulted in reducing the length formed by the branching program;

— the use of the divided sets to protect against attacks of mixed products and mixed input;

— clarifying a number of procedures carried out at different stages of indistinguishable code obfuscation.

УДК 004.94

И.И. Лившиц

ФОРМИРОВАНИЕ КОНЦЕПЦИИ МГНОВЕННЫХ АУДИТОВ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

Лившиц И.И. **Формирование концепции мгновенных аудитов информационной безопасности.**

Аннотация. В данной публикации рассмотрена проблема формирования концепции мгновенных аудитов информационной безопасности (ИБ), направленной, в т.ч. на обеспечение защиты от угроз «нулевого дня» (“zero-day”). Отмечается, что эффективное противодействие угрозам «нулевого дня» основывается на реализации комплекса упреждающих мер, а не только на внедрении новых технических средств защиты. Ключевой особенностью концепции мгновенных аудитов ИБ является формирование оценки как предела слева уровня защищенности в процессе выполнения аудитов ИБ. Методической базой концепции мгновенных аудитов является семейство стандартов ISO серии 27001 и 19011, дополненное множеством (расширяемым) метрик ИБ для формирования количественной оценки уровня защищенности объекта. Полученные результаты могут найти применение при создании моделей и методов обеспечения аудитов ИБ и непрерывного повышения уровня защищенности объектов, находящихся под воздействием угроз нарушения ИБ.

Ключевые слова: информационная безопасность; система менеджмента информационной безопасности; аудит; менеджмент рисков; угрозы; уязвимости; стандарты.

Livshits I.I. **Formation of the instantaneous Information Security Audit Concept.**

Abstract. This publication discusses the problem concerning the concept of the instantaneous information security (IT-Security) audits directed, including providing protection against “zero-day” threats. It is noted that effective "zero-day" counteraction based on implementation a set of preventive IT-Security controls, but not limited new technical facilities installation only. A key feature of this concept of instantaneous IT-Security audits is to assess how the left limit of the protection level in the process of IT-Security audits performing. Methodological basis of the concept of instantaneous IT-Security audits is ISO 27001 and 19011 standards series, supplemented by many (expandable) IT-Security metrics to quantify the object protection level. The obtained results can find application in create of models and methods of IT-Security audits performing and continuous improvement for object protection under the influence of IT-Security violation threats.

Keywords: Information security; Information Security Management System; audit; risk management; threats; vulnerabilities; Standards.

1. Введение. Проблема выполнения аудитов (как процесс оценки) для больших и/или сложных систем рассматривалась в классических трудах Н. Винера, Р. Кини, Х. Райфа, И. Пригожина [1–4]. В работе Н. Винера отмечено требование невмешательства человека в процесс, начиная с момента ввода исходных данных и до получения результата ([1], стр. 47). В работе Р. Кини и Х. Райфа важное внимание уделено потоку данных, поступающему уже непосредственно в самом процессе. Отмечается, что выработка и

анализ возможных альтернатив действий становится явно зависимым от информации, которая станет известна уже в процессе ([2], стр. 24). В работе И. Пригожина отмечается подход Карла Рубино (*Rubino C.)*, который обращает внимание на философский принцип выполнения любой деятельности, в том числе оценки – при рассмотрении любого предмета не следует стремиться к большей точности, чем допускает природа предмета [3]. Эти постулаты могут быть эффективно применены при решении актуальных проблем в области информационной безопасности (ИБ). В настоящее время представлены различные материалы по актуальной проблеме противодействия угрозам «нулевого дня» (“zero-day”). В частности отмечается, что «любые процессы, управляемые людьми, ненадёжны», поэтому крупнейшие поставщики средств ИБ предлагают «единственный» вариант – только постоянное совершенствование технических средств защиты информации (СрЗИ), в частности, Check Point Threat Emulation и Qualys Continuous Monitoring [5–7]. Подобная оценка представляется коммерчески выгодной, но весьма далекой от решения хорошо известной технической проблемы – противостояния СрЗИ как «брони» и угроз – как «снаряда».

Очевидно, что «гонка вооружения» между целевыми (таргетированными) атаками (“advanced persistent threats”, APT) не приведет в ближайшем времени к повышению уровня защищенности объектов, и это отмечается многими экспертами [8–10]. В этой ситуации предлагается применять не только технический подход (СрЗИ) для противодействия угрозам «нулевого дня», но предложить комбинированный метод, основанный на концепции мгновенных аудитов ИБ. Методической базой концепции мгновенных аудитов является семейство стандартов ISO серии 27001 и 19011, дополненное множеством (расширяемым) метрик ИБ для формирования количественной оценки уровня защищенности объекта [11 – 15]. В частности, рекомендуется применять дополнительно количественные метрики обеспечения ИБ из стандартов ISO серии 20000 (для управления ИТ-услугами, например – SLA) [16] и ISO серии 22301 (для управления непрерывностью бизнеса, например – RTO, RPO) [17]. Для успешного решения отраслевых задач ИБ необходимо принять во внимание дополнительно специфические отраслевые стандарты, в частности для аэродромных комплексов – IATA [18].

Необходимо отметить, что сам процесс аудитов (в том числе ИБ) хорошо известен и является обязательным требованием всех упомянутых стандартов ISO (в Российской Федерации они приняты

как ГОСТ Р ИСО), при этом на усмотрение организации отдаются вопросы планирования (частоты) выполнения аудитов и области охвата (“scope”) [19 – 21]. Именно на процесс аудитов, управляемый по частоте, возлагается задача оперативного (в режиме близком к режиму реального времени) выявления уязвимостей в информационных системах (ИС), которые могут быть использованы при реализации угроз «нулевого дня». В стандарте ISO 19011 установлены требования по формированию объема программы аудита, фокусирования на вопросах, наиболее важных для организации, принятие во внимание событий ИБ, произошедших утечек и иных инцидентов [15]. Для формирования концепции мгновенных аудитов ИБ, как средства противодействия АТР, представляется полезным применить известное математическое понятие предела функции, точнее, предела слева, которое позволит формировать количественные оценки защищенности в процессе выполнения аудитов ИБ.

2. Постановка задачи. Как отмечалось выше, в настоящее время для решения проблемы противодействия угрозам «нулевого дня» предлагается «единственный» вариант – только постоянное совершенствование технических СрЗИ, оснащенных новыми («виртуальными», «сканирующими», «аналитическими» и пр.) модулями, способными противостоять АРТ [8 – 10]. В тоже время не приходится ожидать, что процесс постоянного совершенствования только технических СрЗИ приведет к видимому успеху, т.к. охватывает только некоторую часть (технических уязвимостей) инфраструктуры безопасности. В частности, методология систем менеджмента информационной безопасности (СМИБ) рассматривает значительно больше уровней иерархии защиты и типов объектов (в терминологии ISO – “asset”), соответственно, предлагается и значительно больше мер (средств) обеспечения ИБ (в терминологии ISO – “control”) [12]. Более того, расширение перечня применяемых стандартов ISO позволит реализовать интегрированную систему безопасности для выбранных критичных объектов, когда СМИБ дополняется требованиями указанных выше стандартов ISO [16, 17]. В равной мере в указанных стандартах ISO отражено и требование выполнения аудитов и требование обеспечения безопасности, которые могут быть реализованы как в рамках отдельной СМИБ, так и интегрированной системы безопасности [11, 16, 17].

Реализация данных требований в предлагаемой концепции дополняется еще одним важным параметром – требуемой частотой выполнения аудитов с целью максимального повышения

осведомленности и скорости принятия адекватных решений об уязвимостях, которые могут быть использованы злоумышленниками для реализации АТР, об объективной оценке текущего уровня обеспечения ИБ. В этих условиях постановка задачи формулируется следующим образом – разработка концепции мгновенных аудитов ИБ на методической базе риск-ориентированных стандартов ISO, с целью обеспечения комплексного подхода для оценивания защищенности ценных для бизнеса объектов с любой требуемой частотой.

3. Обоснование практической ценности мгновенных аудитов. Практическая ценность предлагаемой концепции мгновенных аудитов основана на известных фактах, что порядка 96% успешных взломов можно было бы избежать, если бы был внедрен ряд простых мер ИБ, а более 75% атак использовали уже известные уязвимости, которые могли бы быть «закрыты» регулярными патчами безопасности [7, 8]. При этом отмечается, что 85% реально произошедших вторжений были обнаружены спустя месяцы (среднее время обнаружения – 5 месяцев) [7, 9].

Дополнительно представляется целесообразным отметить отчет ЦБ РФ за 2014 г. с актуальными данными по оценке обеспечения ИБ на уровне пользователей [22]. В целом банковская система РФ продемонстрировала способность останавливать от 46% до 38% несанкционированных операций (НСО), а средняя сумма одной НСО составляет 335 тыс. руб. В большинстве случаев НСО, связанные с попытками списания денежных средств посредством систем дистанционного банковского обслуживания, произошли вследствие воздействия вредоносного кода на используемое устройство. Также распространенной причиной НСО являлось применение социальной инженерии с использованием сети «интернет», электронной почты и услуг, предоставляемых операторами связи (распространение информации, побуждающей клиента сообщать информацию, необходимую для осуществления переводов денежных средств, в т.ч. информацию аутентификации).

В качестве мер противодействия угрозам «нулевого дня» в настоящее время применяются различные подходы, направленные, в основном, на пресечение последствий потенциально возможных угроз, но не на выявление и устранение уязвимостей, например:

1. «Песочницы», имитирующие рабочие станции организации, в которых анализируются запускаемые файлы на предмет возможных деструктивных воздействий;

2. Анализ аномальной сетевой активности, который осуществляется путем сравнения текущей сетевой активности с построенной эталонной моделью сетевого поведения;

3. Поведенческий анализ рабочих станций, основанный на сравнении активности рабочих станций с эталонной моделью (на уровне самой рабочей станции).

Соответственно, для атак «нулевого дня» (реакция на которые крайне критична по времени) указанные выше примеры дают известный эффект только при постоянном наращивании вычислительных ресурсов для сокращения времени «аналитических» проверок СрЗИ в режиме, близком к режиму реального времени. При этом не инициируется объективный анализ всей совокупности потенциальных уязвимостей и не затрагивается уровень технологических, программных и иных уязвимостей [11, 14].

Рассмотрим дополнительно обоснование адекватности результатов оценки уровня защищенности для ИС, получаемых в случае применения концепции мгновенных аудитов. В работе Ф. Перегудова и Ф. Тарасенко отмечается, что адекватность подразумевает выполнение определенных требований «не вообще», а в той мере, которая достаточна для достижения цели. Можно дать оценку адекватности, если ввести количественную меру, или, цитируя точно: “количественно выражаемую меру адекватности” ([4], стр. 51). Также можно применить рекомендации Р. Кини и Х. Райфа по введению групповых решений, которые, цитируя точно: “систематизируют решение конкретных проблем” ([2], стр. 22). На практике эти рекомендации применяются для формирования количественных метрик, пригодных, в том числе, для групповых оценок деятельности в области ИБ, например – динамика количества выявленных уязвимостей в ИС по результатам аудитов ИБ.

Процесс аудитов [15], как любой процесс оценки, предполагает получение объективных оценок на основании свидетельств аудита, которые затем могут быть воспроизведены, и дополнительно проверены независимыми экспертами (в соответствии с критериями аудита). При этом сам процесс оценки также предполагает определенные временные рамки (как было показано выше [2]), при этом управление «частотностью аудита» позволяет более оперативно контролировать динамику процесса изменения уровня защищенности против любых изменений (соответственно – «динамической перестройки» критериев аудита) [19, 20]. Важным преимуществом предложенной концепции является акцентирование именно на

получении численных оценок, а не простого «соответствия» или «несоответствия». Именно периодическое систематическое получение измеримых численных оценок ИБ, представляется практически полезным для лиц, принимающих решение (ЛПР). Соответственно, адекватность результатов оценки уровня защищенности ИС допустимо трактовать, во-первых, как соответствие установленным критериям аудита, во-вторых, соответствие процессным требованиям аудита ИБ, в-третьих – получение «текущих» значений уровня реализации мер (средств) обеспечения ИБ, необходимых для поддержки принятия «разумных решений» ЛПР [2].

4. Требования ISO к проведению аудитов СМИБ. Требования выполнения аудитов СМИБ на постоянной циклической основе определены в стандарте ISO/IEC 17021:2006. В частности, отмечается, что «программа аудита должна включать в себя проведение двухэтапного первичного аудита, надзорных аудитов в течение первого и второго года и ресертификационного аудита – в течение третьего года до истечения срока действия сертификата. Трехлетний цикл сертификации начинается с принятия решения о сертификации или ресертификации» (п. 9.1.1). Также отмечается частота выполнения надзорных аудитов – «надзорные аудиты (инспекционный контроль) должны проводиться, по крайней мере, один раз в год. Проведение первого надзорного аудита (инспекционного контроля) с момента первоначальной сертификации должно быть не позже, чем через 12 мес. после последнего дня второго этапа аудита» (п. 9.3.2.2). Аналогичные требования предъявляются к проведению аудитов по требованию PCI DSS, в частности все три вида аудитов (QSA – внешний, ISA – внутренний и SAQ – самооценка) проводятся с периодичность 1 раз в год.

Последовательность выполнения аудитов СМИБ с учетом требований ISO/IEC 17021:2006 состоит из: CA – сертификационных аудитов (1-й и 2-й этапы соответственно), SA – надзорных аудитов (1-го и 2-го года соответственно) и RA – ресертификационного аудита (см. рисунок 1). Для данной публикации важно, что интервалы ($t_1 - t_0$), ($t_2 - t_1$) и ($t_3 - t_2$) в общем случае (без экстраординарных ситуаций) равны. Дополнительно необходимо отметить, что аналогичного подхода придерживаются и иные международные системы аудитов, в частности – IATA [18, 23, 24]. В стандарте ISO 19011 (п. 5.2, e), h), j) непосредственно указано, что цели аудита должны формироваться с учетом: правовых и иных других требований, которые организация принимает на себя;

показателей деятельности организации (случаи возникновения нарушений, инцидентов или жалоб потребителей) и результатов предыдущих аудитов [15].

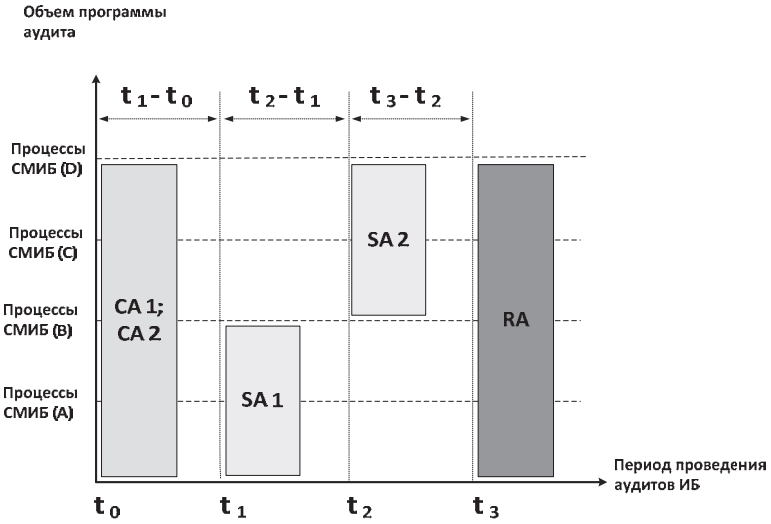


Рис. 1. Схема выполнения аудитов СМИБ

Также ISO 19011 (п. 5.3.3) при формировании объема программы аудита рекомендует принять во внимание факторы: законодательные, контрактные и другие требования, которые организация обязана выполнять; результаты предыдущих внутренних или внешних аудитов; существенные изменения в организации (ее деятельности); возникновение событий внутреннего и внешнего характеров, таких как утечки секретной информации, действия преступного характера [15]. Соответственно, представляется нелогичным и экономически нецелесообразным постоянно осуществлять значительные затраты на применение только дорогостоящих СрЗИ – если, например, на уровне рабочих станций не выполняются требования доменных политик ИБ, на уровне пользователей – не выполняется информирование о правилах работы в сети интернет, на уровне руководителей – не выполняются аудиты ИБ, анализ отчетов и принятия безотлагательных мер в области ИБ.

5. Концепция мгновенных аудитов СМИБ. Концепция мгновенных аудитов предполагает реализацию принципа выполнения аудитов ИБ с частотой, определяемой высшим менеджментом (ЛПР) и

зависящей от предыдущего состояния «слева» уровня защищенности объекта [15, 18, 24 – 26]. Иными словами, если предыдущий Аудит_1 ИБ, проведенный, предположим, месяц назад (отметка t_0) выявил ряд несоответствий (в терминах [15, 18]) и показал, что 40% компьютеров по-прежнему работают под Windows XP с SP2, на 60% рабочих станций пользователи обладают правами администратора, на 70% ноутбуков обновление антивируса не выполняются и/или отключены, то оценка (отметка t_1) текущего уровня защищенности $R_{base} \mid t_1 <= R_{base} \mid t_0$, т.е. не выше предыдущей (см. рисунок 2).

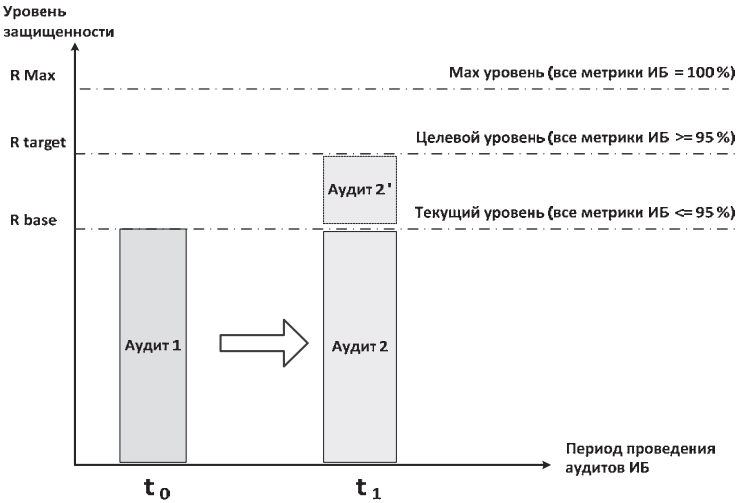


Рис. 2. Оценка достижения уровней защищенности

Также маловероятно и экономически нецелесообразно проводить Аудит_2′ в надежде достигнуть на отметке t_1 целевой уровень защищенности R_{target}, например, 95% (смотри рисунок 2). Соответственно, текущая защищенность объекта (отметка t_1) $R_{target} \mid t_1$ соответствует оценке слева (отметка t_0) $R_{base} \mid t_0$ при отсутствии изменений в состоянии защищенности объекта, выявленных предыдущем Аудит_1. При изменении на интервале ($t_0 - t_1$) состава мер (средств) ИБ, закрытия выявленных на Аудит_1 несоответствий (например, проведения дополнительного обучения), выполнение последующего аудита (Аудит_2′) может иметь смысл для достижения R_{target}. Важно, что частота выполнения аудитов ИБ определяется, в

том числе и допустимым уменьшением интервала ($t_0 - t_1$), например, с ежегодного (как это принято в СМИБ, PCI DSS, IATA) до ежемесячного (еженедельного) и чаще – по требованию ЛПР.

Проблема определения оптимальной частоты аудитов ИБ определяется решением ЛПР на основании полученных наборов оценок защищенности и проведенного анализа в рамках стандартной процедуры «Анализ со стороны руководства» (“Management review”) [11, 15, 17, 18]. Очевидно, что бессмысленно выполнять подряд аудиты ИБ друг за другом, не успевая исправить выявленные несоответствия, не успевая полностью реализовать комплекс корректирующих мер. В частности, метрикой для «старта» следующего аудита ИБ, может являться скорость «замыкания» мини-цикла PDCA, которая, объективно, формирует предел $\lim (t_k - t_i)$. Соответственно, для достижения R_{target} период аудитов ИБ может уменьшаться как $\lim (t_k - t_i) \rightarrow 0$ (см. рисунок 3).

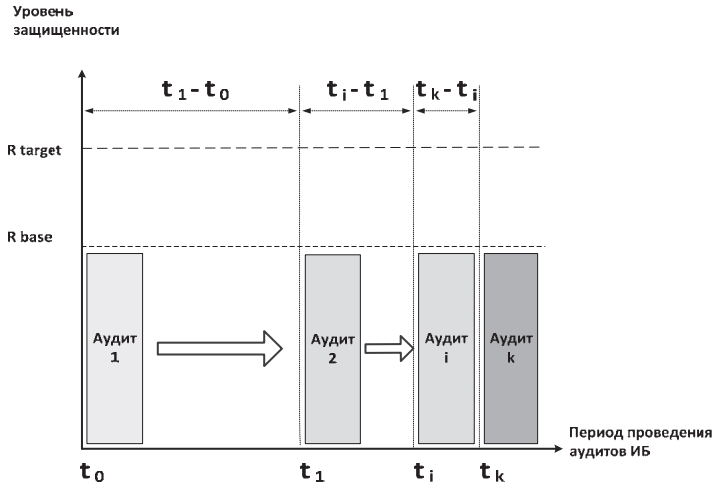


Рис. 3. Снижение периода оперативного противодействия угрозам

Кроме того, для эффективного противодействия APT необходимо уменьшить период выявления, анализа и «закрытия» несоответствий, так как успешная реализация этого процесса представляется значительно быстрее, чем выбор, закупка, доставка, установка и настройка новых и новых СрЗИ. При этом, во-первых, выполняются все требования ISO, во-вторых, дополнительно выполняются требования ЛПР (не желающих ждать целый год для приведения ИБ к требуемому бизнесом уровню защищенности), а в-

третьих, решаются вопросы оперативного противодействия современным угрозам (в пределе «время реакции» СМИБ ($t_k - t_i$) стремится к нулю).

6. Система метрик процесса мгновенных аудитов. Любая концепция обеспечения ИБ требует обоснования у ЛПР (владельцев ценных для бизнеса активов) численных оценок достигнутого (реального) уровня защищенности. Предлагаемая концепция мгновенных аудитов применяет систему количественных (численных) метрик ИБ на базе:

1. Рекомендации SANS;
2. Рекомендации PCI DSS (например, версии 3.0);
3. Стандарты ISO (например, ISO 27004);
4. Стандарты Центрального банк РФ (например, СТО БР ИББС);
5. Документы ФСТЭК (например, приказ ФСТЭК № 31).

Рассмотрим на примере систему мер (средств) обеспечения ИБ, сформированную на базе рекомендаций SANS [6, 9], документов ФСТЭК и стандарта ISO 27001 [11]. Примем во внимание, что предлагаемый состав метрик ИБ не является фиксированным (возможно расширение по требованию ЛПР, регуляторов, контрагентов и пр.), и рекомендуемые метрики содержат не только СрЗИ, но и комплекс организационных мер (см. таблица 1).

Таблица 1. Соответствие мер (средств) обеспечения ИБ

№ п.п.	Контроль SANS	Мера ИБ (ISO 27001)	Мера защиты информации (ФСТЭК)
1.	Учет авторизованных (неавторизованных) устройств	A.8.1.1 – A.8.1.4, A.11.2.8	ИАФ.2
2.	Учет авторизованного (неавторизованного) ПО	A.8.1.1 – A.8.1.4	ИАФ.7 ОПС.0 - ОПС.4
3.	Безопасная конфигурация рабочих станций, серверов, ноутбуков	A.12.1, A.18.2.3	ЗСВ.7, ЗИС.29 УКФ.0 – УКФ.5
4.	Постоянное обнаружение и оценка уязвимостей	A.12.6, A.17.1.2	АНЗ.1 ОБР.1
5.	Защита почтовых приложений	A.12.5.1, A. 13.2.3	АВЗ.0 – АВЗ.3
6.	Прикладное ПО для обеспечения ИБ	A.9.4.4	ИАФ.7 ЗИС.25
7.	Контроль беспроводных соединений	A.9.1.2	УПД.14 ЗИС.3, ЗИС.20
8.	Резервирование (архивирование) данных	A.12.3.1	ОДТ.2 ЗСВ.8, ДНС.4
9.	Обучение и тренинги в области ИБ	A.7.2.2	ДНС.2 ИПО.0 – ИПО.3
10.	Безопасная конфигурация сетевых устройств	A.13.1.1	УПД.3

7. Обоснование математической базы концепции мгновенных аудитов. Для формирования оценки защищенности по результатам аудитов ИБ необходимо применять достоверные математические понятия, дающие обоснование предложенной концепции, в частности одностороннего предела (точнее, предела функции слева). Число $A \in R$ называется левым пределом (или пределом слева) функции *f(x)* в точке a, если для всякого положительного числа ε отыщется отвечающее ему положительное число δ, такое, что для всех точек x из интервала (a – δ, a) справедливо неравенство [27]:

$$| f(x) - A | < \varepsilon.$$

или

$$\lim_{x \to a-0} f(x) = A \Leftrightarrow \forall \varepsilon > 0 \; \exists \delta = \\ = \delta(\varepsilon) > 0 \; \forall x \in (a - \delta, a): | f(x) - A | < \varepsilon.$$

Производная функции *f (x)*:

$$\lim_{\Delta x \to 0} = \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim \frac{d}{dx} f(x) = f'(x).$$

Соответствующий односторонний предел называют левой производной, обозначают $f'_{-}(x)$ [27, 28].

8. Пример определения частных производных для мгновенных аудитов ИБ. Левая производная позволяет оценить требуемый интервал, на котором допустимо (по времени) могут быть выполнены необходимые изменения в СМИБ и обосновано проведение нового аудита ИБ. Для цели противодействиям угрозам «нулевого дня» рассмотрим действительную функцию переменных:

$$y = f(x_1, x_2, x_3, \dots, x_n),$$

где, например, первые 4 переменные описывают атрибуты аудитов ИБ:

x_1 – частота проведения аудитов, определяемая как отношение кол-ва аудитов в СМИБ к наблюдаемому периоду;

x_2 – объем программы аудитов, определяемый как отношение кол-ва охваченных процессов к общему кол-ву процессов в заявленной области сертификации СМИБ;

x_3 – метрика достижения уровня защищенности, определяемая как мера результативности СМИБ $R_{\mathbf{base}} / R_{\mathbf{Max}}$;

x_4 – метрика выполнения корректирующих действий, запланированных на интервал проведения аудитов ИБ.

Тогда частная производная первого порядка по первой переменной x_1 имеет вид:

$$\lim_{\Delta x_1 \to 0} = \frac{f(x_1 + \Delta x_1, x_2, x_3, \ldots, x_k) - f(x_1, x_2, x_3, \ldots, x_k)}{\Delta x_1} = \frac{\partial}{\partial x_1} f(x).$$

Для одной изменяемой переменной x_1 (например, частоты проведения аудитов ИБ) оценим практическое значение частной производной (при неизменности иных переменных), получаем оценку скорости роста уровня защищенности СМИБ:

$$\frac{\partial}{\partial x_1} = f'x_1(x_1, x_2, x_3, \ldots, x_n) = \frac{\Delta R_k}{\Delta t k}.$$

Реализация концепции мгновенных аудитов для оценки защищенности ценных для бизнеса активов с любой требуемой частотой, может быть продемонстрирована как сокращение периода (увеличение частоты) проведения аудитов ИБ при использовании предела слева функции переменных. Заметим, что предложенная концепция позволяет дополнительно исправлять возможные ошибки, присущие сложному процессу аудита, методом локализации обратным процессом, как показано в работе Н. Винера ([1], стр. 222). «Второй контур контроля», реализующий локализацию ошибки стартует с точки, где она замечена, но крайне важно обеспечить, чтобы проверка и отработка выявленной ошибки шла с такой же скоростью, как и сам процесс аудита ИБ, иначе «эффективная скорость» процесса обеспечения ИБ (в составе СМИБ или ИСМ) может снижаться из-за более медленного процесса аудита ИБ.

Отметим снова, что полная «скорость реакции» СМИБ определяется частотой аудитов ИБ, что значительно превышает скорость полного цикла обновлений даже наилучших отраслевых решений CheckPoint [6, 7, 10]. При этом объективно повышается способность системы (СМИБ или ИСМ) эффективно противодействовать угрозам «нулевого дня» в режиме, близком к режиму реального времени. В примере для одной переменной x_1 продемонстрировано увеличение скорости роста уровня защищенности СМИБ $\frac{\Delta R_k}{\Delta t k}$ при известных переменных процесса аудитов ИБ (см. рисунок 4).

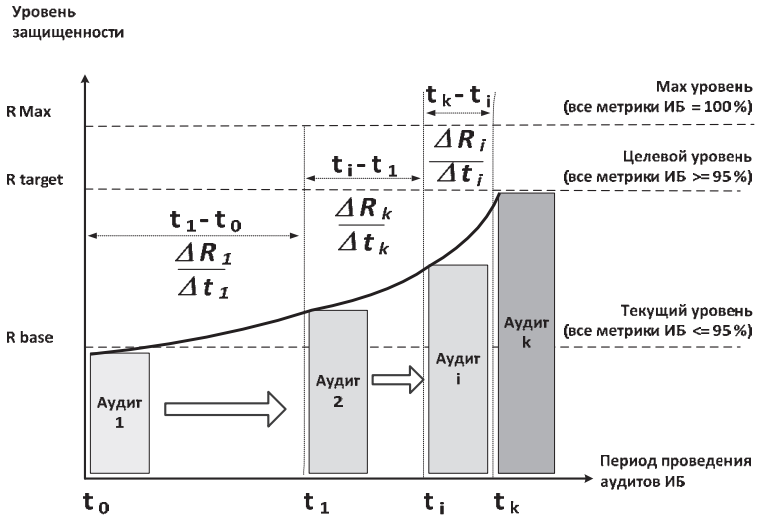


Рис. 4. Пример увеличения скорости роста уровня защищенности

9. Заключение. Предлагаемая концепция мгновенных аудитов ИБ базируется на формировании оценки «предела слева» функции переменных, характеризующих процесс выполнения аудитов ИБ, и направлена на создание непрерывной системы комплексного обеспечения ИБ, в том числе для защиты от угроз «нулевого дня» (“zero-day”).

Литература

1. *Винер Н.* Кибернетика, или управление и связь в животном и машине. 2-е издание // М.: Наука; Главная редакция изданий для зарубежных стран. 1983. 344 с.
2. *Р.Л. Кини, Х. Райфа.* Принятие решений при многих критериях: Предпочтения и замещения: Пер. с англ./ Под ред. И.Ф. Шехнова // М.: Радио и Связь. 1981. 560 с.
3. *Пригожин И., Стенгерс И.* Время. Хаос. Квант. К решению парадокса времени // М.: Едиториал УРСС, 2003. 240 с.
4. *Перегудов Ф.И., Тарасенко Ф.П.* Введение в системный анализ // М.: Высшая школа. 1989. 360 с.
5. Официальный сайт Center for strategic and International Studies. URL: www.csis.org (дата обращения 07.07.2015).
6. Официальный сайт Infosecurity Russia. URL: www.infosecurityrussia.ru (дата обращения 07.07.2015).
7. Официальный сайт Trustwave. URL: www.trustwave.com (дата обращения 07.07.2015).
8. An Osterman Research White Paper «Dealing with Data Breaches and Data Loss Prevention» // Osterman Research, Inc. 2015.

9. Официальный сайт Reuters. URL: www.reuters.com (дата обращения 07.07.2015).
10. *Morvay Z., Gvozdenac D.* Applied Industrial Energy and Environmental Management // John Wiley & Sons, Chichester, UK, 2008.
11. ISO/IEC 27001:2013. Information technology. Security techniques. Information security management systems // Requirements, International Organization for Standardization. 2013. 23 p.
12. ISO/IEC 27000:2014. Information technology. Security techniques. Information security management systems // Overview and vocabulary, International Organization for Standardization. 2014. 31 p.
13. ISO/IEC 27004:2009. Information technology. Security techniques. Information security management systems // Measurement, International Organization for Standardization. 2009. 55p.
14. ISO/IEC 27005-2011 Information technology. Security techniques. Information security management systems // International Organization for Standardization. 2011. 68 p.
15. ISO 19011:2011. Guidelines for auditing management systems // International Organization for Standardization, 2011. 44 p.
16. ISO/IEC 20000-1:2011. Information technology. Service management. Part 1: Service management system requirements // International Organization for Standardization. 2011. 26p.
17. ISO 22301:2012. Societal security. Business continuity management systems // Requirements, International Organization for Standardization. 2012. 24 p.
18. IATA Reference Manual for Audit Programs // Effective, 5-rd Edition. 2014.
19. *Лившиц И.И.* Совместное решение задач аудита информационной безопасности и обеспечения доступности информационных систем на основании требований международных стандартов BSI и ISO // Информатизация и Связь. 2013, Вып. 6. С. 62–67.
20. *Лившиц И.И.* Практические применимые методы оценки систем менеджмента информационной безопасности // Менеджмент качества. 2013. Вып. 1. С. 22–34.
21. *Лившиц И.И.* Подходы к применению модели интегрированной системы менеджмента для проведения аудитов сложных промышленных объектов – аэропортовых комплексов // Труды СПИИРАН. 2014. Вып. 6. С. 72–94.
22. Официальный сайт Центрального Банка РФ URL: www.cbr.ru (дата обращения 07.07.2015).
23. *Шмельова Т.Ф., Сікірда Ю.В., Ассаул О.Ю.* Вплив факторів середовища менеджменту авіапіприемства на безпеку авіаційної діяльності // Технологический аудит и резервы производства. 2015. Т. 2. № 3 (22). С. 17-24.
24. *Нехорошкин Н.И.* Проблемы и возможности информационно-аналитического обеспечения аудита проектов и программ // Вестник АКСОР. 2010. Т. 1. № 12. С. 41-45.
25. *Голощапов А.Н., Рыжов И.В.* Общая характеристика и алгоритм проведения внутреннего аудита системы менеджмента качества организации // Экономика и предпринимательство. 2012. № 5 (28). С. 244-248.
26. *Васильков Ю.В., Гущина Л.С.* Система менеджмента рисков как инструмент управления экономикой предприятия // Методы менеджмента качества. 2012. № 2. С. 10-15.
27. *Ильин В. А., Садовничий В. А., Сендов Бл. Х.* Глава 3. Теория пределов. Математический анализ / Под ред. А. Н. Тихонова. — 3-е изд., перераб. и доп. // М.: Проспект, 2006. Т. 1. 672 с.
28. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров // М.: Наука. 1978. 832 с.

References

1. Viner N. *Kibernetica ili upravlenie i svyaz v zhivotnom i mashine*. [Cybernetics or Control and Communication in the Animal and the Machine]. M.: Nauka, 1983. 344 p. (in Russ).
2. Kini R., Raifa X. *Prinyatie resheni pri mnogih kriteriyah: predpochteniya i zamesheniya*. [Decisions with multiple objectives: Preferences and substitution]. M.: Radio i Svyaz,1981. 240 p. (in Russ).
3. Prigozhin I., Stengers I. Vremya. Haos, Kvant. *K resheniu paradoksa vremeni.* [Time. Chaos. Quantum. To the solution of the paradox of time]. M.: Editorial URSS, 2003. 240 p. (in Russ).
4. Peregudov F., Tarasenko F. Vvedenie v sistemnyi analis. [Introduction to system analysis]. M.: Higher school, 1989. 360 p. (in Russ).
5. Oficial'nyi sait "Center for strategic and International Studies" [Official web site of "Center for strategic and International Studies"]. Available at: www.csis.org (accessed 07.07.2015).
6. Oficial'nyi sait "Infosecurity Russia" [Official web site of "Infosecurity Russia"]. Available at: www.infosecurityrussia.ru (accessed 07.07.2015). (in Russ).
7. Oficial'nyi sait Trustwave [Official web site of Trustwave]. Available at: www.trustwave.com (accessed 07.07.2015).
8. An Osterman Research White Paper «Dealing with Data Breaches and Data Loss Prevention». Osterman Research, Inc. 2015.
9. Oficial'nyi sait Reuters [Official web site of Reuters]. Available at: www.reuters.com (accessed 07.07.2015).
10. Morvay Z., Gvozdenac D. Applied Industrial Energy and Environmental Management. John Wiley & Sons, Chichester, UK. 2008.
11. ISO/IEC 27001:2013. Information technology. Security techniques. Information security management systems. Requirements, International Organization for Standardization. 2013. 23 p.
12. ISO/IEC 27000:2014. Information technology. Security techniques. Information security management systems. Overview and vocabulary, International Organization for Standardization. 2014. 31 p.
13. ISO/IEC 27004:2009. Information technology. Security techniques. Information security management systems. Measurement, International Organization for Standardization. 2009. 55p.
14. ISO/IEC 27005-2011 Information technology. Security techniques. Information security management systems. International Organization for Standardization. 2011. 68 p.
15. ISO 19011:2011. Guidelines for auditing management systems. International Organization for Standardization, 2011. 44 p.
16. ISO/IEC 20000-1:2011. Information technology. Service management. Part 1: Service management system requirements. International Organization for Standardization. 2011. 26p.
17. ISO 22301:2012. Societal security. Business continuity management systems. Requirements, International Organization for Standardization. 2012. 24 p.
18. IATA Reference Manual for Audit Programs. Effective, 5-rd Edition. 2014.
19. Livshitz I. [Joint problem solving information security audit and ensure the availability of information systems based on the requirements of international standards BSI / ISO]. *Informatisatia i Svyaz' – Informatization and Communication.* 2013. vol. 6. pp. 62–67. (In Russ).

20. Livshitz I. [Practical purpose methods for ISMS evaluation]. *Menedzhment kachestva – Quality Management*. 2013. vol. 1. pp. 22–34 (In Russ).
21. Livshitz I. [Approaches to the application of the integrated management system model for carrying out audits for complex industrial facilities – airport complexes]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2014. vol. 6, pp. 72–94. (In Russ).
22. Oficial'nyi sait Cenral'nogo Banka RF [Official web site of Central Bank of RF]. Available at: www.cbr.ru. (accessed 07.07.2015). (In Russ).
23. Shmeleva T., Sikirda Y., Assaul O. [The influence of environmental factors management of the airline safety aviation activity]. *Tehnologicheskij audit i rezervy proizvodstva – Technological audit of production and reserves.* 2015. vol. 2 part 3. pp. 17–24 (in Ukrainian).
24. Nechorochkin N. [Challenges and opportunities of information and analytical support for the projects and programmes audit]. *Vestnik AKCOR – Bulletin AKSOR*. 2010. vol. 1, part 12, pp. 41–45. (In Russ).
25. Goloshapov A., Ryzhov I. [General characteristics of the algorithm and the internal audit of the quality management system of the organization]. *Ekonomika i predprinimatel'stvo – Economics and Business*. 2012. vol. 5. pp. 244–248. (in Russ).
26. Vasil'kov Y., Gushina L. [The risk management system as a tool of enterprise economic management]. Metody menegementa kachestva – Methods of Quality Management. 2012. vol. 2. pp. 10–15. (In Russ).
27. Il'in V., Sadovnichi V., Sendov B. *Glava 3. Teoria predelov. Matematicheskiy analis* [Chapter 3. Theory of limit. Mathematical Analysis]. M.: Prospect. 2006. vol. 1. 672 p. (in Russ).
28. Korn G., Korn T. *Spravochnik po matematike dlya nauchnich rabotnikov I inzhenerov* [Mathematics handbook for scientist and engineers]. M.: Nauka, 1978. 832 p. (in Russ).

Лившиц Илья Иосифович — к-т техн. наук, ведущий аналитик, ООО "Газинформсервис". Область научных интересов: системный анализ, защита информации, риск-менеджмент. Число научных публикаций — 50. Livshitz.il@yandex.ru; 197082, Санкт-Петербург, Богатырский пр.; р.т.: +7(921) 934-48-46.

Livshitz Ilya Iosifovich — Ph.D., lead analyst, LLC "Gasinformservice". Research interests: system analyses, IT-security, risk-management. The number of publications — 50. Livshitz.il@yandex.ru; 197082, Saint-Petersburg, Bogatirskyi str.; office phone: +7(921) 934-48-46.

РЕФЕРАТ

Лившиц И.И. **Формирование концепции мгновенных аудитов информационной безопасности.**

В данной публикации кратко рассмотрена проблема формирования концепции мгновенных аудитов информационной безопасности (ИБ), направленной, в т.ч. на обеспечение защиты от угроз «нулевого дня» (“zero-day”). В настоящее время представлены различные материалы по актуальной проблеме противодействия угрозам «нулевого дня», в частности отмечается, что «любые процессы, управляемые людьми, ненадёжны».

В этой ситуации предлагается применять не только технические методы, но предложить комбинированный метод, основанный на концепции мгновенных аудитов ИБ. Методической базой концепции мгновенных аудитов является семейство стандартов ISO серии 27001 и 19011, дополненное множеством (расширяемым) метрик ИБ для формирования количественной оценки уровня защищенности объекта.

Для формирования оценки защищенности по результатам аудитов ИБ необходимо применять достоверные математические понятия, дающие обоснование предложенной концепции. Решение поставленной задачи – обеспечение комплексного подхода для оценки защищенности ценных для бизнеса активов с любой требуемой частотой, может быть продемонстрировано как сокращение периода (увеличение частоты) проведения аудитов ИБ при использовании предела слева функции переменных. В примере для одной переменной продемонстрировано увеличение скорости роста уровня защищенности СМИБ при известных переменных процесса аудитов ИБ, что позволяет успешно противодействовать угрозам «нулевого дня».

Данные результаты могут найти применение при создании моделей и методов обеспечения аудитов СМИБ и мониторинга состояния объектов, находящихся под воздействием угроз нарушения ИБ, а также при создании моделей и методов оценки защищенности информации объектов СМИБ.

SUMMARY

Livshits I.I. **Formation of the instantaneous Information Security Audit Concept.**

This publication discusses the problem of formation the concept of the instantaneous information security (IT-Security) audits directed, including providing protection against “zero-day” threats. Various recent materials are presented to the actual problem of counter zero-day threats notes that "*any process-driven people, unreliable*". In this situation it is proposed to use not only a technical methods to counter “zero-day” threats, but to offer a combined method based on the concept of instantaneous IT-Security audits. Methodological basis of the concept of instantaneous audits both the ISO 27001 and ISO 19011 standards, which extended with the set of (extensible) metrics for IT-Security formation to quantify the object protection level.

For the formation of IT-Security assessment on the results of IT-Security audits it is necessary to use accurate mathematical concepts, giving the rationale for the proposed concept. The solution of this problem – providing an integrated approach to IT-Security evaluate of valuable business objects with any desired frequency can be shown as a reduction of the period (increase the frequency of IT-Security audits when using the left limit of the function variables. In the example for one variable was demonstrated an increase in the rate of growth of the ISMS level variables with known IT-Security audits process. These results can be used to create models and methods to ensure the ISMS audits and monitoring of objects under the influence of threats to IT-Security violations, as well as the creation of models and methods of estimation of IT-Security facilities ISMS.

РУКОВОДСТВО ДЛЯ АВТОРОВ

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Труды СПИИРАН» http://www.proceedings.spiiras.nw.ru. При регистрации авторам рекомендуется заполнить все предложенные поля данных, так как это значительно ускорит процесс оформления метаданных к новым статьям.

Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше. При подаче материала в редакцию сначала отправляется только статья в формате *.docx. Для обеспечения требований слепого рецензирования при представлении статьи в журнал авторам необходимо удалить персональные данные, содержащиеся в тексте файла и его свойствах.

Объем основного текста – от 5 до 20 страниц включительно. Формат страницы документа – A5 (148 мм ширина, 210 мм высота); ориентация – портретная; все поля – 20 мм. Верхний и нижний колонтитулы страницы – пустые. Основной шрифт документа – Times New Roman, основной кегль (размер) шрифта – 10 pt. Переносы разрешены. Абзацный отступ устанавливается размером в 10 мм. Межстрочный интервал – одинарный. Номера страниц не проставляются.

Не допускается использования цветных шрифтов, цветовых выделений и цветных рисунков. Статьи должны быть полностью готовы к черно-белой печати.

Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материала выделение подразделов.

В основную часть допускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

ISSN 2078-9181

Вып. 6(43)

ТРУДЫ СПИИРАН
proceedings.spiiras.nw.ru

С-Пб
2015